Topics in Optimization and Vector Optimization

by

PETER AU

Thesis

Submitted to the Faculty of the Graduate School of

The Chinese University of Hong Kong

(Division of Mathematics)

In partial fulfillment of the requirements

for the Degree of

Master of Philosophy

Hong Kong

July, 1999

Abstract

Let $A \cdot x \leqslant b$ be a system of linear inequalities, where $A = (a_{ij})$ is an $m \times n$ matrix of real numbers, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. In 1952, Alan J. Hoffman proved that if the system is consistent, then an global error bound exists and depends on A only. Similar result for convex quadratic inequalities system was obtained by Xiao Dong Luo and Zhi Quan Luo when the system is consistent and satisfies the Slater condition. Following the results of Adrian S. Lewis and J. S. Pang, we will present some necessary and sufficient conditions for the existence of an global error bound for an extended real-valued closed proper convex function. Also the works of O. L. Mangasarian and Sien Deng provided some sufficient conditions when considering a system of convex inequalities in stead. Finally, we study the scalarization of Henig proper efficient point, and the relation between Pareto optimizing sequence and scalarly stationary sequence in vector optimization problem.

內容撮要

設 A · x≦b 爲一線性不等式系統，當中 A 是一個 mxn 實數矩陣，$x \in R^n$ 及 $b \in R^m$。在 1952 年，Alan J. Hoffman 証明了當上述系統有解時，就存在了一個整體誤差上限及該上限只依賴於 A。對於凸二元不等式系統，Xiao Dong Luo 和 Zhi Quan Luo 在該系統有解及符合 Slater condition 時，証明了類似結果。跟隨著 Adrian S. Lewis 和 J. S. Pang 的結果，本論文會提出一些充份及必須條件關於誤差上限的存在，當該系統爲一延拓實值封閉真凸函數。另外，O. L. Mangasarian 和 Sien Deng 得出了一些充份條件關於整體誤差上限的存在，當考慮一個凸不等式系統。最後，本論文會討論向量最優化中 Heing 真效率點的純量化及 Pareto 最優列與純量駐列的關係。

ACKNOWLEDGMENTS

I wish to express my gratitude to my supervisor, Professor K. F. Ng. He answered all my questions, even the stupid ones, and was willing to explain patiently throughout the period of my postgraduate studies and in the preparation of this thesis.

Contents

Chapter 1

Introduction

This thesis concerns some recent theoretical results on optimization. Our study is divided into two parts. This first part is devoted to study some error bound results. The second part emphasizes on vector optimization problem. We will study the scalarization of Heing proper efficient points and the relation between Pareto optimizing sequences and scalarly stationary sequences.

The second chapter of this thesis is preliminaries. Some terminology, notations and results on convex analysis and non-smooth analysis are provided for the usage of later chapters. Moreover, well-posedness and exact penalization are introduced.

The main subject in chapter three is to survey various error bounds result on optimization. The study of error bound was first pioneered by Hoffman ([10]). He wanted to know, for a system of linear inequalities, whether there exists an error bound, that is, there exists a constant $c \in \mathbb{R}$ such that $dist(x, S) \leqslant c\|(A \cdot x - b)^+\|$ for all $x \in \mathbb{R}^n$. He proved that the answer is positive whenever the system is consistent in the sense that S is non-empty. Then X. D. Luo and Z. Q. Luo ([14]) extended the result to any consistent polynomial systems. However, besides the constant c, they found that an extra factor $(1 + \|x\|)^{r'}$ should be introduced for the error bound to hold in general. Moreover, when the polynomial system concerned is convex quadratic and satisfies a Slater condition, their work showed us that the extra factor can be removed by the virtue of Theorem 3.2.4. For error bounds of a convex inequality, we will study some necessary and sufficient conditions by Lewis and Pang ([13]). They characterized the existence of an error bound by using normal cones and tangent cones. For example,

$$\text{for all } \bar{x} \in f^{-1}(0) \text{ and } d \in N(\bar{x}; S) \cap T(\bar{x}; dom(f)), f'(\bar{x}; d) \geqslant r^{-1}\|d\|$$

if and only if a global error bound exists. In the last part of this chapter, we will investigate error bound results for convex inequalities systems. Systems satisfying strong slater constraint qualification and slater condition for recession functions will be discussed.

In chapter four, we study the set $H(A, C)$ of all Henig proper efficient points. Characterization theorems (by Zheng [21]) of the set $H(A, C)$ by monotone Minkowski functionals and continuous norms are presented. Therefore, one can translate a vector optimization problem to a scalar optimization problem. In the last part of this chapter, we turn our attention to study Pareto optimizing sequence and scalarly stationary sequence. Theorem of whether a scalarly stationary sequence is weakly Pareto optimizing, asymptotically Pareto optimizing or converges to a weakly Pareto point is given.

Chapter 2

Preliminaries

In this chapter we review some basic results in convex analysis and non-smooth analysis which will be used in the sequal. Our basic references are [1], [2], [3], [6], [8], [9], [12], [18], [19] and [20].

2.1 Recession and Conjugate Functions

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be an extended real-valued function.

Definition 2.1.1 *(a) The effective domain of f is defined to be*

$$dom(f) = \{ x \in \mathbb{R}^n \mid f(x) < +\infty \}.$$

(b) If $dom(f)$ is non-empty, then f is said to be closed.

(c) The epigraph of f is defined to be

$$epi(f) = \{ (x, y) \mid x \in \mathbb{R}^n, y \in \mathbb{R} \text{ and } y \geqslant f(x) \}.$$

(d) If $epi(f)$ is closed in $\mathbb{R}^{n+1}$, then f is said to be closed.

Definition 2.1.2 *f is said to be lower semi-continuous at a point $x \in \mathbb{R}^n$ if*

(a) $x \in dom(f)$ and

(b) for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$f(y) \geqslant f(x) - \epsilon \quad \textit{for all } y \in \mathbb{B}(x, \delta),$$

where $\mathbb{B}(x, \delta)$ denotes the open ball with centre x and radius δ.

Remark: If f is continuous, then f is also semi-continuous.

The following theorem characterizes the lower semi-continuity by level sets and epigraph of f.

Theorem 2.1.1 *(cf. [19, Thm. 7.1]) The following statements are equivalent:*

(a) f is lower semi-continuous on $\mathbb{R}^n$;

(b) $\{ x \in \mathbb{R}^n \mid f(x) \leqslant \alpha \}$ is closed for any $\alpha \in \mathbb{R}$;

(c) f is closed.

Definition 2.1.3 *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be a proper convex function and C be a non-empty convex set in $\mathbb{R}^n$.*

(a) The recession cone C^∞ of C is defined to be

$$C^\infty = \{ y \in \mathbb{R}^n \mid x + \lambda y \in C \textit{ for all } \lambda \geqslant 0 \textit{ and } x \in C \};$$

(b) the recession function f_∞ of f is defined to be

$$epi(f_\infty) = (epi(f))^\infty;$$

(c) the conjugate function $f^\star$ of f is defined to be

$$f^\star(x) = \sup\{ \langle x, y \rangle - f(y) \mid y \in \mathbb{R}^n \} \quad \textit{for } x \in \mathbb{R}^n;$$

(d) the indicator function I_C of C is defined to be

$$I_C(x) = \begin{cases} 0 & \textit{if } x \in C, \\ +\infty & \textit{if } x \notin C; \end{cases}$$

(e) the support function $\delta_C^\star$ of C is defined to be

$$\delta_C^\star(x) = \sup\{\, \langle x, y\rangle \mid y \in C \,\} \quad \textit{for } x \in \mathbb{R}^n.$$

Theorem 2.1.2 *(cf. [19, Thm. 8.1]) Let C be a non-empty convex set in $\mathbb{R}^n$. Then*

(a) C^∞ is a convex cone;

(b) $C^\infty = \{\, y \in \mathbb{R}^n \mid C + y \subseteq C \,\}$.

Theorem 2.1.3 *[19, Thm. 8.5] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\, +\infty \,\}$ be a proper convex function. Then f_∞ is a positively homogeneous proper convex function and*

$$f_\infty(y) = \sup\{\, f(x+y) - f(x) \mid x \in dom(f) \,\} \quad \textit{for all } y \in \mathbb{R}^n.$$

Furthermore, if f is closed, then f_∞ is also closed and for any $x \in dom(f)$, $y \in \mathbb{R}^n$ one has

$$f_\infty(y) = \sup_{\lambda > 0} \frac{f(x+\lambda y) - f(x)}{\lambda} = \lim_{\lambda \to \infty} \frac{f(x+\lambda y) - f(x)}{\lambda}.$$

Theorem 2.1.4 *[19, Thm. 13.3] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\, +\infty \,\}$ be a proper convex function. Then the support function of $dom(f)$ is the recession function $(f^\star)_\infty$ of $f^\star$. Furthermore, if f is closed, then one has the support function of $dom(f^\star)$ is the recession function f_∞ of f.*

Corollary 2.1.1 *[19, Cor. 13.3.4] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\, +\infty \,\}$ be a closed proper convex function. Fix a vector $x^\star \in \mathbb{R}^n$ and define*

$$g(x) = f(x) - \langle x, x^\star\rangle \quad \textit{for } x \in \mathbb{R}^n.$$

Then

(a) $x^\star \in \overline{dom(f^\star)}$ if and only if $g_\infty(y) \geqslant 0$ for all $y \in \mathbb{R}^n$;

(b) $x^\star \in ri(dom(f^\star))$ *if and only if* $g_\infty(y) > 0$ *for all* y *except those satisfying* $-g_\infty(-y) = g_\infty(y) = 0$*;*

(c) $x^\star \in int(dom(f^\star))$ *if and only if* $g_\infty(y) > 0$ *for all* $y \neq 0$*;*

(d) $x^\star \in aff(dom(f^\star))$ *if and only if* $g_\infty(y) = 0$ *for all* y *such that* $-g_\infty(-y) = g_\infty(y)$*,*

where $\overline{dom(f^\star)}$*,* $ri(dom(f^\star))$*,* $int(dom(f^\star))$*,* $aff(dom(f^\star))$ *are the closure, relative interior, interior and affine hull of* $dom(f^\star)$ *respectively.*

2.2 Directional derivative and Subgradient

Definition 2.2.1 *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be any function, and let* $x \in dom(f)$*. The directional derivative of* f *at* x *with respect to a vector* y *is defined to be*

$$f'(x;y) = \lim_{\lambda \downarrow 0} \frac{f(x+\lambda y) - f(x)}{\lambda},$$

if it exists ($+\infty$ *and* $-\infty$ *being allowed).*

Definition 2.2.2 *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be any function.* $x^\star \in \mathbb{R}^n$ *is said to be a subgradient of* f *at a point* x *if*

$$\langle x^\star, z - x \rangle \leqslant f(z) - f(x) \quad \text{for all } z \in \mathbb{R}^n.$$

The set of all subgradients of f *at* x *is called the subdifferential of* f *at* x *and is denoted by* $\partial f(x)$*. When* $\partial f(x)$ *is non-empty,* f *is said to be subdifferentiable at* x*.*

Theorem 2.2.1 *(cf. [19, Thm. 23.1]) Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a proper convex function, and let* $x \in dom(f)$*. For any* $y \in \mathbb{R}^n$*, the difference quotient in Definition 2.2.1 is a non-decreasing function of* $\lambda > 0$*. Therefore* $f'(x;y)$ *exists and*

$$f'(x;y) = \inf_{\lambda > 0} \frac{f(x+\lambda y) - f(x)}{\lambda}.$$

Moreover, $f^{'}(x;y)$ is a positively homogeneous convex function of y and

$$-f^{'}(x;-y) \leqslant f^{'}(x;y) \quad \textit{for all } y \in \mathbb{R}^n.$$

Theorem 2.2.2 *(cf. [19, Thm. 23.2]) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function, and let $x \in dom(f)$. Then $x^\star \in \partial f(x)$ if and only if*

$$f^{'}(x;y) \geqslant \langle x^\star, y\rangle \quad \textit{for all } y \in \mathbb{R}^n.$$

Theorem 2.2.3 *(cf. [19, Thm. 23.4]) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. For $x \notin dom(f)$, $\partial f(x)$ is empty. For $x \in ri(dom(f))$, $\partial f(x)$ is non-empty, $f^{'}(x;y)$ is closed and proper as a function of y, and*

$$f^{'}(x;y) = \sup\{\langle x^\star, y\rangle \mid x^\star \in \partial f(x)\} = \delta^\star_{\partial f(x)}(y).$$

Moreover, $\partial f(x)$ is non-empty and bounded if and only if $x \in int(dom(f))$, in which case $f^{'}(x;y)$ is finite for any $y \in \mathbb{R}^n$.

Theorem 2.2.4 *(cf. [20, Prop. 5A]) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function, and let $x \in dom(f)$. The following statements are equivalent:*

(a) x is the global minimizer of f;

(b) x is a local minimizer of f;

(c) $f^{'}(x;y) \geqslant 0$ for all $y \in \mathbb{R}^n$;

(d) $0 \in \partial f(x)$.

Remark: Theorem 2.2.4 follows easily from Theorem 2.2.2.

Theorem 2.2.5 *(cf. [9, Cor. 1.4.4]) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a closed proper convex function and let x,$s \in \mathbb{R}^n$. Then the following statements are equivalent:*

(a) $f(x) + f^\star(s) - \langle s, x\rangle \leqslant 0$;

(b) $s \in \partial f(x)$;

(c) $x \in \partial f^\star(s)$.

Definition 2.2.3 *Let $C \subseteq \mathbb{R}^n$ be non-empty, closed and convex.*

(a) The distance function of C is defined to be

$$dist(x, C) = \min\{\, \|x - y\| \mid y \in C \,\} \quad \text{for } x \in \mathbb{R}^n,$$

and the unique minimizer is called the projection of x into C and is denoted by $\Pi_C(x)$.

(b) Let $x \in C$. The normal cone of C at x is defined to be

$$N(x; C) = \{\, y \in \mathbb{R}^n \mid \langle c - x, y \rangle \leqslant 0 \quad \text{for all } c \in C \,\}.$$

(c) Let $x \in C$. We say that $d \in \mathbb{R}^n$ is a direction tangent to C at x if there exists a sequence $\{\, x_n \,\} \subseteq C$ and a sequence of positive scalar $\{\, t_n \,\}$ such that

$$\lim_{n\to\infty} x_n = x, \quad \lim_{n\to\infty} t_n = 0, \quad \text{and } \lim_{n\to\infty} \frac{x_n - x}{t_n} = d.$$

The set of all such directions is called the tangent cone of C at x and is denoted by $T(x; C)$.

Alternatively, the normal cone and the tangent cone can be defined in different ways, as the following two theorems show that.

Theorem 2.2.6 *(cf. [8, Prop. 5.2.1]) Let $C \subseteq \mathbb{R}^n$ be non-empty, closed and convex and $x \in C$. Then*

$$T(x; C) = \overline{cone(C - x)}.$$

Theorem 2.2.7 *[8, Prop. 5.2.4 and Cor. 5.2.5] Let $C \subseteq \mathbb{R}^n$ be non-empty, closed and convex. Let $x \in C$. Then the tangent cone of C at x is the polar of the normal cone of C at x and vice versa, that is,*

$$T(x;C) = \{\, d \in \mathbb{R}^n \mid \langle s, d\rangle \leqslant 0 \text{ for all } s \in N(x;C)\,\} \quad \text{and}$$
$$N(x;C) = \{\, d \in \mathbb{R}^n \mid \langle s, d\rangle \leqslant 0 \text{ for all } s \in T(x;C)\,\}.$$

Theorem 2.2.8 *(cf. [8, Prop. 5.3.5]) Let $C \subseteq \mathbb{R}^n$ be non-empty, closed and convex and let $x \in C$. For any $y \in \mathbb{R}^n$, we have*

$$\Pi_C'(x;y) := \lim_{t \downarrow 0} \frac{\Pi_C(x+ty) - \Pi_C(x)}{t} = \Pi_{T(x;C)}(y).$$

The following theorem is evident from definitions.

Theorem 2.2.9 *(cf. [20, Prop. 3J]) Let $C \subseteq \mathbb{R}^n$ be non-empty, closed and convex and let $x \in C$. The subdifferential of the indicator function of C is:*

$$\partial I_C(x) = N(x;C).$$

Theorem 2.2.10 *[8, Ex. 3.3] Let $C \subseteq \mathbb{R}^n$ be non-empty, closed and convex. The subdifferential of $dist(\cdot, C)$ is:*

$$\partial dist(x,C) = \begin{cases} N(x;C) \cap \mathbb{B} & \text{if } x \in C, \\ \{\, \frac{x - \Pi_C(x)}{\|x - \Pi_C(x)\|}\,\} & \text{if } x \notin C, \end{cases}$$

where $\mathbb{B}$ is the closed unit ball in $\mathbb{R}^n$. Hence, $N(x;C) = cone(\partial dist(x,C))$ for any $x \in C$.

Theorem 2.2.11 *(cf. [12, Section 0.3.2]) Let X be a Banach space. For any $x \in X \backslash \{\,0\,\}$, we have*

$$\partial\|x\| = \{\, x^\star \in X^\star \mid \|x^\star\| = 1 \text{ and } \langle x^\star, x\rangle = \|x\|\,\}.$$

Theorem 2.2.12 *(cf. [19, Thm. 23.7]) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\,+\infty\,\}$ be a proper convex function, and let x be a point such that f is subdifferentiable at x but f does not achieve its minimum at x. Then, $N(x; f^{-1}(-\infty, f(x)]) = \overline{cone(\partial f(x))}$.*

Furthermore, if x belongs to the interior of $dom(f)$, then the $cone(\partial dist(x, C))$ is closed automatically. In fact, we have:

Corollary 2.2.1 *(cf. [19, Cor. 23.7.1]) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function, and let $x \in int(dom(f))$ such that $f(x)$ is not the minimum of f. Then,*

$$N(x; f^{-1}(-\infty, f(x)]) = cone(\partial f(x)).$$

Theorem 2.2.13 *[19, Thm. 23.8] Let $f_1, \ldots, f_m$ be proper convex functions on $\mathbb{R}^n$ and let $f = f_1 + \cdots + f_m$. Then*

$$\partial f_1(x) + \cdots + \partial f_m(x) \subseteq \partial f(x) \quad \textit{for all } x \in \mathbb{R}^n.$$

Furthermore, if $\bigcap_{i=1}^m ri(dom(f_i)) \neq \emptyset$, then the above inclusion becomes equality.

Theorem 2.2.14 *[20, Thm. 5C] Let f_1, f_2 be two closed proper convex functions on $\mathbb{R}^n$, and let $x \in dom(f_1) \cap dom(f_2)$. Suppose that $dom(f_1) \cap int(dom(f_2)) \neq \emptyset$. Then*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x).$$

Theorem 2.2.15 *[3, p.234] Let f_1, f_2 be two closed proper convex functions on $\mathbb{R}^n$ and both continuous at a point $x_0 \in \mathbb{R}^n$. Define $g(x) := \max\{ f_1(x), f_2(x) \}$ on $\mathbb{R}^n$. Then*

$$\partial g(x_0) = co(\partial f_1(x_0) \cup \partial f_2(x_0)).$$

The following theorem is a more general version of the above theorem.

Theorem 2.2.16 *[12, Thm. 4.2.2] Let T be a compact topological space and X be a locally convex space. Let $f(t, x)$ be a function from $T \times X$ to $\mathbb{R} \cup \{+\infty\}$ which satisfies*

(a) $f(t, x)$ is convex in x for every $t \in T$,

(b) $f(t,x)$ *is upper semicontinuous in* t *for every* $x \in X$ *and*

(c) $f(t,x)$ *is continuous at* x_0 *for every* $t \in T$.

Set $f_t(x) := f(t,x)$ *on* X, $f(x) := \sup_{t \in T} f(t,x)$ *on* X *and* $T(x_0) := \{t \in T \mid f(t,x_0) = f(x_0)\}$. *Then*

$$\partial f(x_0) = \overline{co\{\, \partial f_t(x_0) \mid t \in T(x_0)\,\}}.$$

Furthermore, if X *is a real reflexive Banach space, then the closure can be dropped.*

Theorem 2.2.17 *[19, Thm. 24.7] Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a closed proper convex function, and let* S *be a non-empty closed bounded subset of* $int(dom(f))$. *Then the set*

$$\partial f(S) = \bigcup_{x \in S} \partial f(x)$$

is non-empty, closed and bounded. The real number

$$\alpha = \sup\{\, \|x^\star\| \mid x^\star \in \partial f(S)\,\} < \infty$$

and has the properties that

$$f'(x;z) \leqslant \alpha \|z\| \qquad \text{for all } x \in S \text{ and for all } z;$$

$$|f(y) - f(z)| \leqslant \alpha \|y - z\| \qquad \text{for all } y \in S \text{ and for all } x \in S.$$

2.3 Well-Posedness and ϵ-subgradient

Now, we introduce the definition of well-posedness.

Definition 2.3.1 *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a closed proper convex function.* f *is said to be well-posed if every stationary sequence of* f *is minimizing. That is, for every sequence* $\{x_n\}$ *in* $\mathbb{R}^n$, *if there exists* $d_n \in \partial f(x_n)$ *for each* n *such that* $\lim_{n\to\infty} d_n = 0$, *then* $\lim_{n\to\infty} f(x_n) = \inf\{\, f(x) \mid x \in \mathbb{R}^n\,\}$.

For any $\epsilon > 0$, $x^\star \in \mathbb{R}^n$ is said to be an ϵ-subgradient of f at x if

$$\langle x^\star, y - x \rangle \leqslant f(y) - f(x) + \epsilon \quad \text{for all } y \in \mathbb{R}^n,$$

and the set of all ϵ-subgradient of f at x, denoted by $\partial_\epsilon f(x)$, is called the ϵ-subdifferential of f at x.

Theorem 2.3.1 *[18, Prop. 3.15] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be a closed proper convex function and $x \in dom(f)$. Then, $\partial_\epsilon f(x) \neq \emptyset$ for any $\epsilon > 0$.*

Definition 2.3.2 *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be a closed proper convex function. A sequence $\{ x_n \} \subseteq \mathbb{R}^n$ is said to be an ϵ-stationary sequence of f if there exist sequences $\{ x_n^\star \} \subseteq \mathbb{R}^n$ and $\{ \epsilon_n \}$ of positive scalars such that*

$$x_n^\star \in \partial_{\epsilon_n} f(x_n) \quad \textit{for each } n, \quad \lim_{n \to \infty} x_n^\star = 0 \quad \textit{and} \lim_{n \to \infty} \epsilon_n = 0.$$

Theorem 2.3.2 *[1, Prop. 2.1] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be a closed proper convex function. f is well-posed if and only if any ϵ-stationary sequence of f is minimizing.*

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be a closed proper convex function. For any $\lambda > m := \inf\{ f(x) \mid x \in \mathbb{R}^n \}$, we define

(a) $r(\lambda) = \inf\{ \|c\| \mid c \in \partial f(x) \text{ and } f(x) = \lambda \}$,

(b) $l(\lambda) = \inf\{ \frac{f(x)-\lambda}{dist(x, f^{-1}(-\infty,\lambda])} \mid x \notin f^{-1}(-\infty, \lambda] \}$,

(c) $k(\lambda) = \inf\{ f'(x; d/\|d\|) \mid d \in \partial f(x) \text{ and } f(x) = \lambda \}$.

Then we can characterize the well-posedness by the three scalars.

Theorem 2.3.3 *[1, Thm. 2.2] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be a closed proper convex function. Suppose $f^{-1}(-\infty, \lambda] \subseteq ri(dom(f))$ for all $\lambda > m$. Then the following statements are equivalent:*

(a) f is well-posed;

(b) $r(\lambda) > 0$ *for all* $\lambda > m$;

(c) $l(\lambda) > 0$ *for all* $\lambda > m$;

(d) $k(\lambda) > 0$ *for all* $\lambda > m$.

Theorem 2.3.4 *[2, Thm. 2.3] Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a closed proper convex function and consider the minimization problem* (P)*:*

$$\min\ f(x)$$
$$s.t.\ x\ \in \mathbb{R}^n.$$

If $0 \in ri(dom(f^\star))$*, then*

(a) (P) *has a non-empty optimal solution set of the form* $S = K + E^\perp$ *where* $E^\perp$ *is the orthogonal complement of aff(dom(*$f^\star$*)) and* K *is a compact subset of* E.

(b) for any ϵ*-stationary sequence* $\{x_n\}$*, we have*

(i) $\{x_n\}$ *is minimizing (i.e.* f *is well-posed.);*

(ii) $\lim_{n\to\infty} dist(x_n, S) = 0$;

(iii) the projected sequence $\{\Pi_E x_n\}$ *is bounded and all its cluster points belong to* $S \cap E = K$.

2.4 Exact Penalization

Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $S \subseteq \mathbb{R}^n$ be non-empty, closed and convex. Consider the following constrained minimization problem:

$$\begin{aligned} &\min\ g(x) \\ &\text{s.t. } x \in S. \end{aligned} \tag{2.1}$$

Definition 2.4.1 *Let S be given as above. $p : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be a penalty function of S if*

$$p(x) = \begin{cases} 0 & \text{if } x \in S, \\ > 0 & \text{if } x \notin S. \end{cases}$$

Definition 2.4.2 (Exact Penalty Property) *Let the constrained minimization problem* (2.1) *have a non-empty solution set. A penalty function p of S is said to have the exact penalty property if there exists an $\alpha \geqslant 0$ such that the following unconstrained minimization problem:*

$$\min \; g(x) + \alpha p(x)$$
$$s.t. \;\; x \in \mathbb{R}^n,$$

has a solution belonging to S.

Definition 2.4.3 *Let $g : \mathbb{R}^n \to \mathbb{R}$. g is said to be Lipschitz of rank $K \geqslant 0$ if*

$$|g(x) - g(y)| \leqslant K\|x - y\| \quad \text{for all } x, y \in \mathbb{R}^n.$$

The following theorem shows that the distance function can play the role of exact penalty function in some cases.

Theorem 2.4.1 *[8, Thm. 1.2.3] Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $S \subseteq \mathbb{R}^n$ be non-empty, closed and convex. For any $\bar{x} \in S$, the following statements are equivalent*

(a) $\bar{x}$ is the global minimizer of g on S;

(b) there exists $\alpha > 0$ such that $\bar{x}$ is the unconstrained global minimizer of the following function on $\mathbb{R}^n$:

$$x \mapsto g(x) + \alpha dist(x, S).$$

We can compute the constant α explicitly whenever g is Lipschitz, as the following theorem shows.

Theorem 2.4.2 *[6, Prop. 2.4.3] Let $g : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz of rank K and S be a non-empty subset of $\mathbb{R}^n$. Suppose that $\bar{x} \in S$ is a global minimizer of g on S. Then for any $K' \geqslant K$, $\bar{x}$ is an unconstrained global minimizer of the following function on $\mathbb{R}^n$:*

$$x \mapsto g(x) + K' dist(x, S).$$

Chapter 3

Some Recent Results on Error Bounds

In the chapter, we survey some results of error bound in optimization. These results are taken from [7], [10], [13], [14] and [16].

3.1 Hoffman's Error Bound

Consider the following system of linear inequalities

$$\begin{cases} A_1 \cdot x = a_{11}x_1 \ + \cdots + a_{1n}x_n \leqslant b_1 \\ \quad\vdots \qquad\qquad\qquad\qquad \vdots \\ A_m \cdot x = a_{m1}x_1 + \cdots + a_{mn}x_n \leqslant b_m \end{cases} \tag{3.1}$$

or simply $A \cdot x \leqslant b$, where $A = (a_{ij})$ is an $m \times n$ matrix of real numbers with row vectors $A_1, \ldots, A_m$, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. We assume that the system is consistent in the sense that the solution set S defined by (3.1) is non-empty. We will present the results in [10] showing that a global error bound exists for S. That is, there exists a constant $c > 0$ such that, for the distance of x to the set S,

$$dist(x, S) \leqslant c \left\| (A \cdot x - b)^+ \right\| \quad \text{for all } x \in \mathbb{R}^n, \tag{3.2}$$

where z^+ denotes the m-vector whose ith coordinate is $z_i^+ = \max\{ z_i, 0 \}$ if $z = (z_1, \ldots, z_m)$. Thus $\|(A \cdot x - b)^+\|$ is a quantity measuring the extent for x to violate the system (3.1) and therefore (3.2) is a condition implying that the distance of x to the solution set of (3.1) is bounded by its extent of failing (3.1). The sketch of the proof for (3.2) was given in [10] with omission of the proof of a few lemmas. Below we present the proofs in full: in particular the idea of proof of Lemma 3.1.3, due to the author, is given in terms of polar cones.

Fix $y \notin S$ and write $\hat{y} = \Pi_S(y)$. Since S is clearly closed and convex, $\hat{y}$ does exist and is unique. Let I be the set of all active indices i for $\hat{y}$: $i \in I$ if and only if $A_i \cdot \hat{y} = b_i$.

Note that,

$$i \notin I \text{ if and only if } A_i \cdot \hat{y} < b_i. \tag{3.3}$$

By continuity, (3.3) implies that when i is inactive, $A_i \cdot z < b_i$ for all z sufficiently near to $\hat{y}$. In particular, since $\hat{y}$ is nearest to y among all elements of S, it follows that I is non-empty (If I were empty then all z near to $\hat{y}$ satisfy (3.1), contradicting the nearest property of $\hat{y}$ to y). Let

$$S_a := \{ x \in \mathbb{R}^n | \quad A_i \cdot x \leqslant b_i \quad \text{for all} \quad i \in I \}.$$

Note for instance that $\hat{y} \in S_a$. We have the following lemma.

Lemma 3.1.1 *$y \notin S_a$ and $\|y - \hat{y}\| = dist(y, S_a)$.*

Proof

Suppose the contrary that $y \in S_a$. Then, by the convexity of S_a, we have the line-segment $(y, \hat{y}) \subset S_a$. By (3.3), we can find a $y_0 \in (y, \hat{y}) \subset S_a$ sufficiently close to $\hat{y}$ such that $A_i \cdot y_0 < b_i$ for all $i \notin I$. Since $y_0 \in S_a$, it follows that $y_0 \in S$ and $\|y - y_0\| < \|y - \hat{y}\|$ contradicting to the fact that $\hat{y}$ is the projection of y in S. Now, it remains to show that $\|y - \hat{y}\| = dist(y, S_a)$. Take $z \in S_a$ be such that

$\|y - z\| = dist(y, S_a)$. We claim that z=$\hat{y}$. To show this, suppose not. Then, the line-segment $(\hat{y}, z)$ is non-degenerate, and each x' in $(\hat{y}, z)$ satisfies the inequality $A_i \cdot x' \leqslant b_i$ for all $i \in I$ as $\hat{y}, z \in S_a$. Moreover (3.3) implies that $A_i \cdot x' < b_i$ for all $i \notin I$ whenever $x' \in (\hat{y}, z)$ sufficiently close to $\hat{y}$. Thus $x' \in S$ for all such x'. Further, in view of the variational inequality (which must be satisfied by the definition of z) ,

$$\langle y - z, \hat{y} - z \rangle \leqslant 0. \tag{3.4}$$

Write $x' = \hat{y} + t(z - \hat{y}) \in S$ for some $t \in (0, 1)$, we have

$$\langle y - x', \hat{y} - x' \rangle < 0 \tag{3.5}$$

because

$$\begin{aligned}
\langle y - x', \hat{y} - x' \rangle &= \langle y - \hat{y} - tz + t\hat{y}, t(\hat{y} - z) \rangle \\
&= \langle y - z + (1 - t)z - (1 - t)\hat{y}, t(\hat{y} - z) \rangle \\
&= t\langle y - z, \hat{y} - z \rangle - (1 - t)t\langle \hat{y} - z, \hat{y} - z \rangle < 0
\end{aligned}$$

by virtue of (3.4), and the fact that $\hat{y} \neq z$. Applying the cosine law

$$\|a\|^2 + \|b\|^2 = \|a + b\|^2 - 2\langle a, b \rangle$$

for vectors a, b in $\mathbb{R}^n$, it follows from (3.5) that

$$\|y - x'\|^2 + \|\hat{y} - x'\|^2 < \|y - \hat{y}\|^2,$$

so $\|y - x'\| < \|y - \hat{y}\|$, contradicting to the fact that $\hat{y}$ is the projection of y in S. □

Definition 3.1.1 *Let F_n be a continuous function from $\mathbb{R}^n$ to $\mathbb{R}$. F_n is said to be*

(a) positively definite if $F_n(x) \geqslant 0$ for all $x \in \mathbb{R}^n$ and the equality holds only when $x = 0$;

(b) positively homogeneous if $F_n(\alpha x) = \alpha F_n(x)$ *for all* $\alpha \geqslant 0$ *and for all* $x \in \mathbb{R}^n$.

Lemma 3.1.2 *Suppose* $F_m : \mathbb{R}^m \to \mathbb{R}$ *satisfies (a) and (b) of the above definition with* $n = m$. *Then there exists an* $e > 0$ *such that for any* $y \in \mathbb{R}^m$ *and any* $J \subseteq \{1, 2, \ldots, m\}$, *the following inequality holds:*

$$F_m(\bar{y}) \leqslant eF_m(y) \quad \textit{whenever} \quad y = (y_1, \ldots, y_m),\ \bar{y} = (\bar{y_1}, \ldots, \bar{y_m})$$

with

$$\bar{y_i} = \begin{cases} y_i, & \textit{if } i \in J \\ 0 & \textit{otherwise.} \end{cases}$$

Proof

When $y = 0$, the result holds obviously. Set

$$e := \max\{ e_J \mid J \subseteq \{1, \ldots, m\} \},$$

where

$$e_J = \sup\{ \frac{F_m(\bar{y})}{F_m(y)} \mid y \in \mathbb{R}^m \backslash \{0\} \} = \sup\{ \frac{F_m(\bar{y})}{F_m(y)} \mid \|y\| = 1 \} < \infty.$$

Here the equality holds because F_m is positively homogeneous and the map $y \mapsto \bar{y}$ is linear; the supremum is attained and is finite because the unit sphere in $\mathbb{R}^m$ is compact, F_m is continuous and positively definite while the map $y \mapsto \bar{y}$ is also continuous.

□

Lemma 3.1.3 *Let* M *be an* $m \times n$ *matrix of real numbers and* $C_M =: \{ z \in \mathbb{R}^n \mid M \cdot z \leqslant 0 \}$. *Let* $E_M =: \{ x \in \mathbb{R}^n \backslash C_M \mid$ *0 is the point in* C_M *nearest to* $x \}$. *Then there exists* $d_M > 0$ *such that*

$$F_m((M \cdot x)^+) \geqslant d_M F_n(x) \quad \textit{for all } x \in E_M, \tag{3.6}$$

where F_m and F_n are given as in Definition 3.1.1.

Remark: *In view of the variational inequality, a non-zero vector x belongs to E_M if and only if $x \cdot z \leqslant 0$ for all $z \in C_M$. Thus E_M is exactly the deleted polar cone of C_M: $E_M = C_M^0 \backslash \{0\}$ where the polar C_M^0 of C_M is, by definition, $\{x \in \mathbb{R}^m \mid x \cdot z \leqslant 0$ for all $z \in C_M\}$.*

Proof

When $x = 0$, (3.6) holds obviously. Let $h(x) = \frac{F_m((M \cdot x)^+)}{F_n(x)}$ for all $x \in \mathbb{R}^n \backslash \{0\}$. Then h is strictly positive on E_M. Indeed, if $x \in E_M$ then $x \notin C_M$, so $(M \cdot x)^+ \neq 0$ and consequently $h(x) > 0$ by positive definiteness of F_m and F_n. Let

$$d := \inf\{h(x) \mid x \in E_M, \|x\| = 1\} = \inf\{h(x) \mid x \in C_M^0, \|x\| = 1\}.$$

Since C_M^0 is clearly closed, it follows from the compactness (of the unit sphere) and the continuity (of h) that $d > 0$ as the infimum is attained. Letting $d_M = d$, (3.6) is seen to hold for all unit vectors $x \in E_M$, and hence for all other vectors in E_M as F_m, F_n are positively homogeneous.

□

Now, we are going to prove the main theorem of this section.

Theorem 3.1.1 (Hoffman's Error Bound [10]) *Let F_n and F_m be given as in Definition* 3.1.1. *Then there exists a constant $c > 0$ such that for any $x \in \mathbb{R}^n$, there exists a solution $\hat{x}$ of* (3.1) *with*

$$F_n(x - \hat{x}) \leqslant cF_m((A \cdot x - b)^+).$$

(*The constant c does not depend on b.*)

Proof

Fix $y \notin S$; let $\hat{y}$ and I be defined as at the beginning. Let M be the $m \times n$ matrix obtained from A by substituting 0 for the ith row when $i \notin I$ and $\bar{b}$ be the vector obtained from b by substituting 0 for the ith components when $i \notin I$. Since I

consists of exactly those i corresponding to which the ith row of the inequality system, $A \cdot \hat{y} \leqslant \bar{b}$ becomes an equality. That is, $M \cdot \hat{y} = \bar{b}$. By Lemma 3.1.1, $y \notin S_a = \{\text{the solution of } M \cdot x \leqslant \bar{b}\}$ and $\|y - \hat{y}\| = dist(y, S_a) > 0$. Note that $S_a - \hat{y} = C_M$ and $\|y - \hat{y}\| = dist(y - \hat{y}, C_M)$, so $y - \hat{y} \in E_M$, where C_M, E_M are as in Lemma 3.1.3. Applying Lemma 3.1.3, we have

$$\begin{aligned} F_m((M \cdot y - \bar{b})^+) &= F_m((M \cdot (y - \hat{y}))^+) \\ &\geqslant d_M F_n(y - \hat{y}) \\ &\geqslant d F_n(y - \hat{y}), \end{aligned}$$

where

$$\begin{aligned} d = \min\{\, d_{M'} \mid d_{M'} &\text{ be defined as in Lemma 3.1.3 and } M' \text{ be a } m \times n \\ &\text{matrix obtained from A by substituting 0 for some rows}\,\} > 0. \end{aligned}$$

Therefore, d depends on A, F_m and F_n only. Further, by Lemma 3.1.2, one has $eF_m((A \cdot y - b)^+) \geqslant F_m((M \cdot y - \bar{b})^+)$. Consequently

$$\begin{aligned} F_n(y - \hat{y}) &\leqslant \frac{1}{d} F_m((M \cdot y - \bar{b})^+) \\ &\leqslant \frac{e}{d} F_m((A \cdot y - b)^+) \end{aligned}$$

Thus $\frac{e}{d}$ has the properties required in the statement of the Theorem 3.1.1.

□

By the above theorem, we have the following corollary easily.

Corollary 3.1.1 *Let A, B be any real $m \times n$ and $j \times n$ matrices respectively. Then there exists a constant $c > 0$ such that whenever the following system*

$$\begin{cases} A \cdot x \leqslant a \\ B \cdot x = b \quad \textit{where } a \in \mathbb{R}^m,\ b \in \mathbb{R}^j \end{cases} \tag{3.7}$$

is consistent, we have for any $x \in \mathbb{R}^n$,

$$dist(x, H) \leqslant c\left(|||(A \cdot x - a)^+||| + \|B \cdot x - b\|\right),$$

where H *is the solution set defined by* (3.7), $|||\cdot|||$, $\|\cdot\|$ *are some fixed norms on* $\mathbb{R}^m$ *and* $\mathbb{R}^j$ *respectively and the distance is with respect to some norm on* $\mathbb{R}^n$.

Proof

Since all the norms on $\mathbb{R}^j$ are equivalent (respectively on $\mathbb{R}^m$), we may assume without loss of generality that $\|\cdot\|$ (respectively $|||\cdot|||$) is the 1-norm. Hence $\|z\| = \|z^+\| + \|z^-\|$ for all $z \in \mathbb{R}^j$, where $z^- = (-z)^+$. Consequently,

$$\|B \cdot x - b\| = \|(B \cdot x - b)^+\| + \|(-B \cdot x + b)^+\|$$

for all $x \in \mathbb{R}^n$. Rewrite (3.7) as

$$\begin{cases} A \cdot x & \leqslant a \\ B \cdot x & \leqslant b \\ -B \cdot x & \leqslant -b \end{cases}$$

By Theorem 3.1.1, for $c > 0$ large enough, we have

$$\begin{aligned} dist(x, H) &\leqslant c\left(|||(A \cdot x - a)^+||| + \|(B \cdot x - b)^+\| + \|(-B \cdot x + b)^+\|\right) \\ &= c\left(|||(A \cdot x - a)^+||| + \|B \cdot x - b\|\right). \end{aligned}$$

□

3.2 Extension of Hoffman's Error Bound to Polynomial Systems

Consider the following system of inequalities:

$$f_1(x) \leqslant 0, f_2(x) \leqslant 0, \ldots, f_m(x) \leqslant 0, g_1(x) = 0, g_2(x) = 0, \ldots, g_k(x) = 0, \quad (3.8)$$

where $x \in \mathbb{R}^n$ and all $f_i^{'}s$, $g_i^{'}s$ are differentiable. Let S denote the solution set of (3.8) and we assume throughout that the system is consistent in the sense that S is non-empty. Let f be the vector-valued function whose ith component function is f_i; let g be defined likewise. In the special case when f and g are some affine linear mappings, we have the Hoffman's error bound result in Corollary 3.1.1. This error bound result cannot be extended to the case where $f_i^{'}s$, $g_i^{'}s$ are arbitrary polynomial mappings as the following counter example shows that there may exist no constant $c > 0$ satisfying the inequality

$$dist(x, S) \leqslant c\,(\|(f(x))_+\| + \|g(x)\|) \quad \text{for all } x \in \mathbb{R}^n. \tag{3.9}$$

Example 1

Let $f : \mathbb{R} \to \mathbb{R}$ be the function $x \mapsto x^2$. Then the solution set S of the inequality system $f(x) \leqslant 0$ is exactly $\{0\}$. Further, for any $x \in \mathbb{R}$, $dist(x, S) = |x|$. For any $c > 0$, we can choose $x \neq 0$ close to zero sufficiently such that $dist(x, S) = |x| > c\,\|(f(x))_+\| = cx^2$. Therefore, the Hoffman's error bound (3.9) fails to hold. Instead, we have $dist(x, S) \leqslant \|(f(x))_+\|^{\frac{1}{2}}$ for all $x \in \mathbb{R}$.

The above example suggests that a possible generalization of Hoffman's error bound to polynomial system is:

$$dist(x, S) \leqslant c\,(\|(f(x))_+\| + \|g(x)\|)^r \quad \text{for all } x \in \mathbb{R}^n, \tag{3.10}$$

where c and r are some positive constants depending on the coefficients and the degrees of the polynomials $f_1, \ldots, f_m$, $g_1, \ldots, g_k$ only. However, (3.10) only holds in some special cases, but not in general (see Example 2 below).

3.2.1 An Error Bound to Polynomial Systems

In this section, we show that in order to make (3.10) holds in general, we need an extra factor $(1+\|x\|)^{r'}$, where r' is a positive constant. To show this, we need the following result of Hormander, the proof is referred to [11] .

Theorem 3.2.1 (Hormander [11]) *Let $f:\mathbb{R}^n \to \mathbb{R}$ be a real polynomial. Let $S := \{\, x \in \mathbb{R}^n \mid f(x) = 0 \,\}$ and assume it is non-empty. Then there exist positive constants c, r and a (possibly negative) constant r' such that*

$$dist(x, S) \leqslant c(1+\|x\|)^{r'}\, |f(x)|^r \quad \text{for all } x \in \mathbb{R}^n.$$

By introducing slack variables z_i for $i = 1, \ldots, m$, Theorem 3.2.1 is extended to a system of polynomial equations and inequalities.

Theorem 3.2.2 (X. D. Luo and Z. Q. Luo [14]) *Let*

$$S := \{\, x \in \mathbb{R}^n \mid f_1(x) \leqslant 0, \ldots, f_m(x) \leqslant 0, g_1(x) = 0, \ldots, g_k(x) = 0 \,\},$$

where $f_i's$, $g_i's$ are polynomials with real coefficients. Suppose that S is non-empty. Then there exists constants $c > 0$, $r > 0$ and $r' \geqslant 0$ such that

$$dist(x, S) \leqslant c(1+\|x\|)^{r'} (\|(f(x))_+\| + \|g(x)\|)^r \quad \text{for all } x \in \mathbb{R}^n.$$

Proof

Consider the polynomial $h : \mathbb{R}^{n+m} \to \mathbb{R}$ given by

$$h(x, z) = (f_1(x) + {z_1}^2)^2 + \cdots + (f_m(x) + {z_m}^2)^2 + {g_1}^2(x) + \cdots + {g_k}^2(x),$$

where $z = (z_1, \ldots, z_m)^T$. Let $\hat{S} := \{\, (x, z) \in \mathbb{R}^{n+m} \mid h(x, z) = 0 \,\}$. Then it is obvious that $x \in S$ if and only if $(x, z) \in \hat{S}$ with $z_i = \sqrt{(-f_i(x))_+}$, where $i = 1, \ldots, m$. By supposition S is non-empty and so is $\hat{S}$. By Theorem 3.2.1, there exists some constants $\alpha > 0$, $\beta > 0$ and $\beta' > 0$ such that

$$dist((x, z), \hat{S}) \leqslant \alpha(1 + \|(x, z)\|)^{\beta'}\, |h(x, z)|^{\beta}$$

for all $(x,z) \in \mathbb{R}^{n+m}$. For any $x \in \mathbb{R}^n$ and let $z = (\sqrt{(-f_i(x))_+}, \ldots, \sqrt{(-f_m(x))_+})$. Take $(\hat{x}, \hat{z}) \in \hat{S}$ be such that $\|(x,z) - (\hat{x},\hat{z})\| = dist((x,z),\hat{S})$. Note that $\hat{x} \in S$ and $a_+ = a + (-a)_+$ for any real number a. Then

$$\begin{aligned}
dist(x,S) &\leqslant \|x - \hat{x}\| \\
&\leqslant \|(x,z) - (\hat{x},\hat{z})\| \\
&= dist((x,z),\hat{S}) \\
&\leqslant \alpha(1 + \|(x,z)\|)^{\beta'} \, |h(x,z)|^{\beta} \\
&\leqslant \alpha(1 + \|x\| + \|z\|)^{\beta'} [(f_1(x)_+)^2 + \cdots + (f_m(x)_+)^2 \\
&\qquad\qquad\qquad\qquad + {g_1}^2(x) + \cdots + {g_k}^2(x)]^{\beta} \\
&\leqslant c(1 + \|x\|)^{r'} (\|(f(x))_+\|^2 + \|g(x)\|^2)^{\beta} \qquad\qquad (\star) \\
&\leqslant c(1 + \|x\|)^{r'} (\|(f(x))_+\| + \|g(x)\|)^{r}, \text{ where } r = 2\beta.
\end{aligned}$$

$(\star)$ is due to the fact that $z_i^2 \leqslant |f_i(x)| \leqslant d(1 + \|x\|)^p$, $i = 1, \ldots, m$, for some positive constants d and p as $f_i's$ are polynomials. Next set c, r' large enough.

□

The above error bound is weaker than that of Hoffman because of the extra factor $(1 + \|x\|)^{r'}$ and the exponent r. One way to remove the extra factor is to restrict the error bound to a bounded set.

Corollary 3.2.1 *Let $f_1, \ldots, f_m, g_1, \ldots, g_k$ and S be defined as in Theorem* 3.2.2. *Let ρ be a positive constant. Then the following inequality holds:*

$$dist(x,S) \leqslant c(\|(f(x))_+\| + \|g(x)\|)^r \quad \textit{for all } x \in B(0,\rho),$$

for some positive constant c and $r > 0$.

The following example shows that we cannot remove the extra factor $(1 + \|x\|)^{r'}$ in general.

Example 2

Consider the following system:

$$-x_1 \leqslant 0, \quad -x_2 + 1 \leqslant 0, \quad x_1 x_2 = 0.$$

Then $S = \{\,(0,x_2) \mid x_2 \geqslant 1\,\}$. For any $t \geqslant 0$, $dist((t,0),S) \geqslant t \to \infty$ as $t \to \infty$. But $\|(f(x))_+\| + \|g(x)\| = \|(-t,1)_+\| + \|0\| = 1$ for all $x = (t,0)$ with $t \geqslant 0$. Therefore, the error bound in Theorem 3.2.2 fails without the extra factor.

We will see in the following section that for convex quadratic inequalities systems satisfying the Slater condition, $(1+\|x\|)^{r'}$ can be removed and the exponent $r = 1$.

3.2.2 Error Bound for Convex Quadratic Inequalities Systems

Consider the following convex quadratic inequalities system:

$$f_1 \leqslant 0, \quad f_2 \leqslant 0 \quad ,\ldots, \quad f_m \leqslant 0, \tag{3.11}$$

where $f_i(x) = \frac{1}{2}\langle x, Q^i \cdot x\rangle + \langle b^i, x\rangle - c_i$, $i = 1, 2, \ldots, m$ with each Q^i an $n \times n$ symmetric positive semi-definite matrix, each b^i an n-vector, each c_i is a real number. We use N(respectively L) to denote the set of all indices i for which each f_i is non-linear(respectively linear) and use S to denote the solution set of (3.11) and assume it is non-empty throughout the following.

Assumption 1(Slater condition)

There exists some $x^\star \in S$ such that $f_i(x^\star) < 0$ for all $i \in N$. Such a point is called a Slater point.

By the above assumption, the following theorem shows that there exists an analogue of Hoffman's error bound for (3.11).

Theorem 3.2.3 (X. D. Luo and Z. Q. Luo [14]) *Suppose that Assumption 1 holds. Then there exists a positive constant c depending on f only such that*

$$dist(x,S) \leqslant c\,\|(f(x))_+\| \quad \text{for all } x \in \mathbb{R}^n,$$

where $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))^T$.

Before proving the above theorem, we need the following definition, theorem and some lemmas.

Definition 3.2.1 (Asymptotic Constraint Qualification (ACQ)) *Let f be a differentiable convex function from $\mathbb{R}^n$ to $\mathbb{R}^m$ with component functions $f_1, \ldots, f_m$. f is said to be satisfying the (ACQ) if there exists a positive constant ρ such that for each $I \subseteq \{1, 2, \ldots, m\}$ and each $x \in S$ with the properties that*

(a) $\{\nabla f_i(x) \mid i \in I\}$ *linearly independent* and

(b) $f_I(x) = 0$,

it holds true that

$$\rho \geqslant \sup\{ \|\lambda_I\| \mid \lambda_I \in \mathbb{R}^{|I|}, \lambda_I > 0 \text{ and } \|\sum_{i \in I} \lambda_i \nabla f_i(x)\| = 1 \}.$$

Theorem 3.2.4 (Mangasarian [15]) *Let f be a differentiable convex function from $\mathbb{R}^n$ to $\mathbb{R}^m$. Let N, L and S be defined as before. Suppose there exists a Slater point $x^\star \in S$ and that f satisfies the (ACQ). Then*

$$dist(x, S) \leqslant \rho\sqrt{n}\, \|(f(x))_+\| \quad \text{for all } x \in \mathbb{R}^n,$$

where ρ is given as in Definition 3.2.1.

Unfortunately, the existence of ρ is not easy to verify in practice. But for the system defined by (3.11), we will show that ρ always exists. Hence, Theorem 3.2.3 is seen to hold as a consequence of Theorem 3.2.4. Therefore, in the following, we will prove that the set $\{ \|\lambda_I\| \}$ is uniformly bounded for all pair of (I, x), where λ_I and (I, x) are defined as above.

Lemma 3.2.1 *For any $n \times n$ symmetric positive semi-definite matrix Q, there exist two positive constants u and v such that*

$$u\,\|Q \cdot x\|^2 \leqslant \langle x, Q \cdot x\rangle \leqslant v\,\|Q \cdot x\|^2 \quad \text{for all } x \in \mathbb{R}^n.$$

Proof

By diagonalization, write $Q = A^T \Lambda A$ where Λ is the diagonal matrix with non-negative eigenvalues, $\lambda_i's$, of Q as its diagonal and A is the matrix obtained from the unit eigenvectors of Q. Therefore, $A^T = A^{-1}$. Write $A \cdot x = (y_1, \ldots, y_n)$. We see that

$$\langle x, Q \cdot x \rangle = \langle x, A^T \Lambda A \cdot x \rangle = \langle A \cdot x, \Lambda A \cdot x \rangle = \sum_{i=1}^{n} \lambda_i y_i^2$$

and

$$\begin{aligned} \|Q \cdot x\|^2 &= \langle A^T \Lambda A \cdot x, A^T \Lambda A \cdot x \rangle = x^T \cdot A^T \Lambda^T A A^T \Lambda A \cdot x \\ &= \langle A \cdot x, \Lambda^2 A \cdot x \rangle = \sum_{i=1}^{n} \lambda_i^2 y_i^2. \end{aligned}$$

Note that the lemma is trivial if all $\lambda_i's = 0$. Now, Suppose $\lambda_i's > 0$ for some i. Then set $u = 1/\max\{\, \lambda_i \mid \lambda_i > 0 \,\}$ and $v = 1/\min\{\, \lambda_i \mid \lambda_i > 0 \,\}$.

$\square$

Lemma 3.2.2 *Let $Q^1, Q^2, \ldots, Q^k$ be any real $n \times n$ matrices. Then there exists $M > 0$ such that for any $x \in \mathbb{R}^n$, there exists an $y \in \mathbb{R}^n$ with the following properties:*

(a) $\|y\| \leqslant M \left(\|Q^1 \cdot y\| + \cdots + \|Q^k \cdot y\| \right)$ *and*

(b) $Q^1 \cdot y = Q^1 \cdot x, \ldots, Q^k \cdot y = Q^k \cdot x.$

Proof

By Hoffman's error bound in Corollary 3.1.1, there exists $M > 0$ depending on $Q^1, \ldots, Q^k$ only such that with any $x \in \mathbb{R}^n$ and the following equality system:

$$\begin{cases} Q^1 \cdot z = Q^1 \cdot x \\ \qquad \vdots \\ Q^k \cdot z = Q^k \cdot x & \text{for } z \in \mathbb{R}^n, \end{cases}$$

one has

$$dist(z, H) \leqslant M \sum_{i=1}^{k} \left\| Q^i \cdot z - Q^i \cdot x \right\| \quad \text{for all } z \in \mathbb{R}^n,$$

where $H := \{ w \in \mathbb{R}^n \mid Q^i \cdot w - Q^i \cdot x = 0 \quad \text{for all } i = 1, \dots, k \}$.

Then we put $z = 0$ and obtain

$$dist(0, H) = \|0 - y\| \leqslant M \sum_{i=1}^{k} \left\| 0 - Q^i \cdot x \right\| = M \left(\left\| Q^1 \cdot y \right\| + \cdots + \left\| Q^k \cdot y \right\| \right)$$

for some $y \in H$. Since $y \in H$, we clearly have $Q^1 \cdot y = Q^1 \cdot x, \dots, Q^k \cdot y = Q^k \cdot x$.
□

Lemma 3.2.3 *Suppose that $Q^1, \dots, Q^k$ are $n \times n$ symmetric positive semi-definite matrices. Let θ be a positive constant. Then, for all $\lambda_1, \dots, \lambda_k$ with $\lambda_i \geqslant \theta$, i=1,...,k, we have the following inequality:*

$$\left\| \lambda_1 Q^1 \cdot x + \cdots + \lambda_k Q^k \cdot x \right\|^2 \geqslant \tau \left(\left\| Q^1 \cdot x \right\|^2 + \cdots + \left\| Q^k \cdot x \right\|^2 \right) \quad \textit{for all } x \in \mathbb{R}^n,$$

where $\tau > 0$ depending on θ and $Q^1, \dots, Q^k$ only.

Proof

Fix $x \in \mathbb{R}^n$ and let y be given as in Lemma 3.2.2. By Lemma 3.2.1, there exists an $u > 0$ depending on $Q^1, \dots, Q^k$ only such that $u \|Q^i \cdot y\|^2 \leqslant \langle y, Q^i \cdot y \rangle$, for $i = 1, \dots k$. Since $\theta \leqslant \lambda_i$ for all i, it follows from Schwarz inequality and Lemma 3.2.2 that:

$$\begin{aligned} \theta u \left\| Q^i \cdot y \right\|^2 &\leqslant \theta \langle y, Q^i \cdot y \rangle \\ &\leqslant \langle y, (\lambda_1 Q^1 + \cdots \lambda_k Q^k) \cdot y \rangle \\ &\leqslant \|y\| \left\| (\lambda_1 Q^1 + \cdots \lambda_k Q^k) \cdot y \right\| \\ &\leqslant M \left(\left\| Q^1 \cdot y \right\| + \cdots + \left\| Q^k \cdot y \right\| \right) \left\| (\lambda_1 Q^1 + \cdots \lambda_k Q^k) \cdot y \right\|. \end{aligned}$$

Taking summation on both sides and making use of the inequality:

$$(1, \dots, 1)(\alpha_1, \dots \alpha_k) \leqslant k^{\frac{1}{2}} (\sum \alpha_i^2)^{\frac{1}{2}},$$

we have

$$\begin{aligned}\theta u \sum_{i=1}^{k} \left\|Q^i \cdot y\right\|^2 &\leqslant Mk \left(\left\|Q^1 \cdot y\right\| + \cdots + \left\|Q^k \cdot y\right\|\right) \left\|(\lambda_1 Q^1 + \cdots \lambda_k Q^k) \cdot y\right\| \\ &\leqslant Mk^{\frac{3}{2}} \left(\left\|Q^1 \cdot y\right\|^2 + \cdots + \left\|Q^k \cdot y\right\|^2\right)^{\frac{1}{2}} \left\|(\lambda_1 Q^1 + \cdots \lambda_k Q^k) \cdot y\right\|.\end{aligned}$$

Dividing both sides by $(\|Q^1 \cdot y\|^2 + \cdots + \left\|Q^k \cdot y\right\|^2)^{\frac{1}{2}}$, we have

$$\theta u \left(\sum_{i=1}^{k} \left\|Q^i \cdot y\right\|^2\right)^{\frac{1}{2}} \leqslant Mk^{\frac{3}{2}} \left\|(\lambda_1 Q^1 + \cdots \lambda_k Q^k) \cdot y\right\|$$

and hence

$$\left\|\lambda_1 Q^1 \cdot x + \cdots + \lambda_k Q^k \cdot x\right\|^2 \geqslant \tau \left(\left\|Q^1 \cdot x\right\|^2 + \cdots + \left\|Q^k \cdot x\right\|^2\right) \quad \text{for all } x \in \mathbb{R}^n,$$

where $\tau = \frac{\theta^2 u^2}{M^2 K^3} > 0$ depending on θ and $Q^1, \ldots, Q^k$ only.

□

Lemma 3.2.4 *Let f_i be given as in (3.11), for $i = 1, \ldots, m$. Let $\{ x^r \}_r$ be a sequence in $\mathbb{R}^n$ such that $f_i(x^r) = 0$, for all i, for all r and further suppose that the sequences $\{ Q^i \cdot x^r \}_r$ and $\{ \langle b^i, x^r \rangle \}_r$ are bounded for all i. Then, there exists a bounded sequence $\{ \bar{x}^r \}_r \subset \mathbb{R}^n$ with the same properties stated above for $\{ x^r \}_r$.*

Proof

We prove the lemma for $m = 1$ first and then generalize the result to any m. Write f, Q, b for f_1, Q^1, b^1 respectively. Since $\mathbb{R}^n = KerQ \oplus (kerQ)^{\perp}$, there exist $y^r \in KerQ$ and $z^r \in (KerQ)^{\perp}$ such that $x^r = y^r + z^r$. Suppose $b \notin (KerQ)^{\perp}$ and use $\bar{b}$ to denote the projection of b to $KerQ$. Let $a^r := y^r - \langle y^r, e\rangle e$, where $e = \bar{b}/\|\bar{b}\|$. We claim that

(a) $Q \cdot a^r = 0$ for all $r \in \mathbb{N}$,

(b) $\langle b, a^r \rangle$=0 for all $r \in \mathbb{N}$,

(c) $\{ x^r - a^r \}_r$ is bounded.

(Note that (a), (b) imply that $f(x^r - a^r) = f(x^r) = 0$ and hence it follows from (c) that the sequence $\{x^r - a^r\}_r$ has the properties required for $\{\bar{x}^r\}_r$, proving the lemma for the case when $m = 1$ and $b \notin (KerQ)^{\perp}$.)

Since $y^r, e \in KerQ$, we have $Q \cdot a^r = 0$ for all $r \in \mathbb{N}$ and so (a) holds.

To show (b), we note that $b - \bar{b} \perp KerQ$. It follows that $\langle y^r, b\rangle = \langle y^r, \bar{b}\rangle$ and $\langle e, b\rangle = \langle e, \bar{b}\rangle$. Therefore,

$$\begin{aligned}\langle b, a^r\rangle &= \langle b, y^r\rangle - \langle y^r, e\rangle\langle b, e\rangle \\ &= \langle \bar{b}, y^r\rangle - \langle y^r, \bar{b}/\|\bar{b}\|\rangle\langle \bar{b}, \bar{b}/\|\bar{b}\|\rangle \\ &= 0.\end{aligned}$$

Hence, (b) holds.

It remains to verify (c). Since $\{Q \cdot x^r\}_r$ is bounded, so is $\{\langle x^r, Q \cdot x^r\rangle\}_r$ by Lemma 3.2.1. On the other hand, by the symmetry of Q, we have

$$\begin{aligned}\langle x^r, Q \cdot x^r\rangle &= \langle y^r + z^r, Q \cdot z^r\rangle \\ &= \langle Q \cdot (y^r + z^r), z^r\rangle \\ &= \langle z^r, Q \cdot z^r\rangle \quad \text{for all } x \in \mathbb{R}^n.\end{aligned}$$

Therefore, $\{\langle z^r, Q \cdot z^r\rangle\}_r$ is also bounded. By Lemma 3.2.1 again, we get that $\{z^r\}_r$ is bounded. Furthermore, since $\{\langle b, x^r\rangle\}_r$ is bounded by assumption and $x^r = y^r + z^r$, it follows that $\{\langle b, y^r\rangle\}_r$ is bounded. Therefore, $\{\langle y^r, e\rangle\}_r$ is bounded since $e = b/\|b\|$. As $x^r - a^r = \langle y^r, e\rangle e + z^r$, we prove (c).

If $b \in (KerQ)^{\perp}$, then (a), (b), (c) hold with $a^r := y^r$. Hence, Lemma 3.2.4 is true for $m = 1$.

Now, we extend the result to $m > 1$. Let $X_1 := KerQ^1 \cap Ker(b^1)$. By the above argument, we can take a sequence $\{a_1^r\}_r \subset X_1$ with $\{x^r - a_1^r\}_r$ bounded. As X^1 is a subspace of $\mathbb{R}^n$, we can restrict f_2 on X^1 and apply the above argument to a_1^r instead of x^r. Then, we get a sequence $\{a_2^r\}_r \subset X^1 \cap KerQ^2 \cap Ker(b^2)$ with $\{a_1^r - a_2^r\}_r$ bounded. Thus, $\{x^r - a_2^r\}_r$ is bounded and $f_i(x^r - a_2^r) = 0$ for

$i = 1, 2$. Inductively, the lemma is true.

□

Lemma 3.2.5 *Suppose that Assumption 1 holds. Then, there exists a positive constant δ depending on $f_i^{'}s$ only such that*

$$\left\| \sum_{i \in N \cup L} \lambda_i \nabla f_i(x) \right\| = \left\| \sum_{i \in N} \lambda_i (Q^i \cdot x + b^i) + \sum_{j \in L} \lambda_j b^j \right\| \geqslant \delta \left(\sum_{i \in N} \lambda_i + \left\| \sum_{j \in L} \lambda_j b^j \right\| \right)$$

for all $\lambda \geqslant 0$ and for all $x \in \mathbb{R}^n$ with $f_i(x) = 0$, $i = 1, \dots, m$, where N, L are the index sets for the non-linear, linear constrains respectively.

<u>**Proof**</u>

We prove this lemma by induction on the cardinality of N.

Obviously, it is true for $|N| = 0$. Suppose the lemma is true for $|N| \leqslant k$. Then we will show that it is also true for $|N| = k + 1$.

By translation to origin if necessary, by Assumption 1, we can suppose that 0 is a Slater point: $0 \in S$ with $f_i(0) < 0$ for all $i \in N$. By definition of f_i, it follows that $c_i > 0$ for $i \in N$ and $c_j \geqslant 0$ for $j \in L$. Suppose on the contrary that the lemma fails for $|N| = k+1$. Then, we can take sequences $\{ x^r \}_r$ and $\{ \lambda^r \}_r$ with $f_i(x^r) = 0$, $\lambda_i^r \geqslant 0$ for each i and each r such that

$$\left\| \sum_{i \in N} \lambda_i^r (Q^i \cdot x^r + b^i) + \sum_{j \in L} \lambda_j^r b^j \right\| \to 0 \quad \text{as } r \to \infty, \tag{3.12}$$

and

$$\sum_{i \in N} \lambda_i^r + \left\| \sum_{j \in L} \lambda_j^r b^j \right\| = 1 \quad \text{for all } r. \tag{3.13}$$

For each r, consider

$$Z^r := \left\{ \mu = (\mu_j)_{j \in L} \mid -\mu_j \leqslant 0 \text{ for all } j \in L \text{ and } \sum_{j \in L} \mu_j b^j = \sum_{j \in L} \lambda_j^r b^j \right\}$$
$$= \{ \mu \mid A \cdot \mu \leqslant 0 \text{ and } B \cdot \mu = a^r \},$$

where $A = -I_{|L|\times|L|}$, $B \cdot \mu = \sum_{j\in L} \mu_j b^j$ and $a^r = \sum_{j\in L} \lambda_j^r b^j$. So A, B do not depend on r. By the Hoffman's error bound: there exists an $c > 0$ (depending on A and B only) such that

$$dist(0, Z^r) \leqslant c\,(\|(A\cdot 0)_+\| + \|B\cdot 0 - a^r\|)\,.$$

Therefore, we may take an $\mu^r \in Z^r$ such that

$$\|\mu^r\| \leqslant c\|a^r\| = c\|\sum_{j\in L} \lambda_j^r b^j\| \leqslant c,$$

where the last inequality is due to (3.13). Hence, replacing λ_j^r by μ_j^r if necessary, we assume that $\{\,\lambda_j^r\,\}_r$ is bounded for all $j \in L$. Obviously, by (3.13) again, $\{\,\lambda_i^r\,\}_r$ is also bounded for all $i \in N$. Therefore, by choosing a subsequence if necessary, we assume that $\{\,\lambda_i^r\,\}_r$ converges to $\lambda_i^\infty \geqslant 0$ for all $i \in N \cup L$. Let $N_1 := \{\, i \in N \mid \lambda_i^\infty > 0\,\}$ and $N_0 := N\backslash N_1$. We claim that $N_1 \neq N$. To show this, suppose on the contrary that there exists an $\theta > 0$ such that $\lambda_i^r \geqslant \theta$ for all $i \in N$ and all r. By (3.12) and (3.13), we know that $\{\,\|\sum_{i\in N}\lambda_i^r Q^i\cdot x^r\|\,\}_r$ is bounded and it follows from Lemma 3.2.3 that $\{\,\|Q^i\cdot x^r\|\,\}_r$ is bounded for all $i \in N$. Since $f_i(x^r) = 0$ for all $i \in N \cup L$, we have the following inequalities by Lemma 3.2.1

$$\left|\langle b^i, x^r\rangle\right| \leqslant c_i + \frac{1}{2}\left|\langle x^r, Q^i\cdot x^r\rangle\right| \leqslant c_i + v\|Q^i\cdot x^r\|^2 \quad \text{for all } r \in N,$$

where the constant $v > 0$ does not depend on r. Therefore, $\{\,\langle b^i, x^r\rangle\,\}_r$ is bounded for all $i \in N \cup L$. (Obviously, $\{\,\langle b^j, x^r\rangle\,\}_r$ is bounded by c_j for $j \in L$.)
By Lemma 3.2.4, we can assume $\{\,x^r\,\}_r$ is bounded with limit x^∞ (considering a subsequence if necessary). Then

$$f_i(x^\infty) = 0 \text{ for } i \in N \cup L \text{ and } \sum_{i\in N}\lambda_i^\infty(Q^i\cdot x^\infty + b^i) + \sum_{j\in L}\lambda_j^\infty b^j = 0,$$

where the second equality is due to (3.12). Let $f(x) = \sum_{i\in N\cup L}\lambda_i^\infty f_i(x)$. Then, $f(x^\infty) = 0$, $\nabla f(x^\infty) = 0$ and also the Slater condition implies that $f(0) < 0$ since

$\lambda_i^\infty > 0$ for $i \in N_1 = N$. By the convexity of f, this is impossible. Thus, N_1 must be a proper subset of N.

Next, we claim that there exists $\alpha > 0$ and $r_0 \in \mathbb{N}$ such that

$$\lambda_i^r \|Q^i x^r\| \geqslant \alpha \quad \text{for all } r \geqslant r_0, i \in N_0. \tag{3.14}$$

Suppose (3.14) does not hold. Then, there exists $l \in N_0$ such that 0 is a cluster point of $\{\, \|\lambda_l^r Q^l x^r\| \,\}_r$. Without loss of generality, assume $\|\lambda_l^r Q^l x^r\| \to 0$ as $r \to \infty$. Hence by (3.12) and the fact that $\lambda_l^r \to 0$, we have

$$\left\| \sum_{i \in N \setminus \{\, l \,\}} \lambda_i^r (Q^i \cdot x^r + b^i) + \sum_{j \in L} \lambda_j^r b^j \right\| \to 0 \quad \text{as } r \to \infty$$

and

$$\sum_{i \in N \setminus \{\, l \,\}} \lambda_i^r + \left\| \sum_{j \in L} \lambda_j^r b^j \right\| \to 1 \quad \text{as } r \to \infty.$$

Therefore, the above two relations contradict to the induction hypothesis for $N \setminus \{\, l \,\}$ whose cardinality is k. Hence, (3.14) holds, and we may therefore assume henceforth that $r_0 = 1$. Consider the linear system in $y \in \mathbb{R}^n$,

$$\begin{cases} Q^i \cdot y = Q^i \cdot x^r & i \in N, \\ \langle b^j, y \rangle = \langle b^j, x^r \rangle & j \in N_1 \cup L. \end{cases} \tag{3.15}$$

Then, by Lemma 3.2.2, the above system has a solution y^r with the property that

$$\|y^r\| \leqslant M \left(\sum_{i \in N} \|Q^i \cdot y^r\| + \sum_{j \in N_1 \cup L} \left| \langle b^j, y^r \rangle \right| \right), \tag{3.16}$$

where M depending on $Q^{i'}s$ and $b^{j'}s$ only.

By Lemma 3.2.1, there exists $u > 0$ depending on $Q^{i'}s$ only such that

$$\langle Q^i \cdot y^r, y^r \rangle \geqslant u \|Q^i \cdot y^r\|^2 \quad \text{for all } r \text{ and } i. \tag{3.17}$$

It follows from (3.14) that for each $i \in N_0$ and $r \geqslant r_0$,

$$\begin{aligned}\lambda_i^r \langle Q^i \cdot y^r, y^r \rangle &\geqslant u\lambda_i^r \|Q^i \cdot y^r\|^2 \\ &= u\lambda_i^r \|Q^i \cdot x^r\|^2 \\ &\geqslant u\alpha \|Q^i \cdot x^r\| \\ &= u\alpha \|Q^i \cdot y^r\|. \end{aligned} \tag{3.18}$$

Since y^r is a solution of (3.15), we have

$$\langle y^r, Q^i \cdot y^r \rangle = \langle y^r, Q^i \cdot x^r \rangle = \langle Q^i \cdot y^r, x^r \rangle = \langle Q^i \cdot x^r, x^r \rangle \quad \text{for all } i \in N;$$

consequently $f_i(y^r) = f_i(x^r) = 0$ for $i \in N_1 \cup L$. Therefore,

$$\langle b^i, y^r \rangle = c_i \quad \text{for } i \in L \text{ and } \left|\langle b^i, y^r \rangle\right| \leqslant c_i + \frac{1}{2}\langle y^r, Q^i \cdot y^r \rangle \quad \text{for } i \in N_1. \tag{3.19}$$

From, (3.16), (3.17), (3.18) and (3.19), we have

$$\begin{aligned}\|y^r\| &\leqslant M\left(\sum_{i\in N_0} \|Q^i \cdot y^r\| + \sum_{i \in N_1} \|Q^i \cdot y^r\| + \sum_{j \in N_1} \left|\langle b^j, y^r\rangle\right| + \sum_{j\in L} \left|\langle b^j, y^r \rangle\right|\right) \\ &\leqslant M(\sum_{i\in N_0} \frac{\lambda_i^r \langle Q^i \cdot y^r, y^r\rangle}{u\alpha} + \sum_{i\in N_1} \frac{\sqrt{\langle Q^i \cdot y^r, y^r\rangle}}{\sqrt{u}} \\ &\qquad + \sum_{j \in N_1 \cup L} c_j + \sum_{j\in N_1} \frac{\langle y^r, Q^j \cdot y^r\rangle}{2}). \end{aligned} \tag{3.20}$$

On the other hand, by using the fact that,

1. $\langle Q^i \cdot y^r + b^i, y^r \rangle = c_i + \frac{1}{2}\langle Q^i \cdot y^r, y^r\rangle \geqslant c_i > 0 \quad$ for $i \in N_1$,

2. $\langle b^j, y^r \rangle = \langle b^j, x^r \rangle = c_j \quad$ for $j \in L$,

3. $Q^i \cdot y^r = Q^i \cdot x^r \quad$ for $i \in N$,

we have, by computing the inner product of $y^r/\|y^r\|$ with the expression in (3.12),

$$\begin{aligned}
&\sum_{i\in N}\lambda_i^r\langle Q^i\cdot x^r+b^i,\frac{y^r}{\|y^r\|}\rangle+\sum_{j\in L}\lambda_j^r\langle b^j,\frac{y^r}{\|y^r\|}\rangle \quad (\to 0 \text{ as } r\to\infty)\\
=&\sum_{i\in N}\lambda_i^r\langle Q^i\cdot y^r+b^i,\frac{y^r}{\|y^r\|}\rangle+\sum_{j\in L}\lambda_j^r\frac{c_j}{\|y^r\|}\\
=&\sum_{i\in N_1}\lambda_i^r\frac{c_i}{\|y^r\|}+\sum_{i\in N_1}\frac{\lambda_i^r}{2}\frac{\langle Q^i\cdot y^r,y^r\rangle}{\|y^r\|}\\
&\quad+\sum_{i\in N_0}\lambda_i^r\frac{\langle Q^i\cdot y^r,y^r\rangle}{\|y^r\|}+\sum_{i\in N_0}\lambda_i^r\frac{\langle b^i,y^r\rangle}{\|y^r\|}+\sum_{j\in L}\lambda_j^r\frac{c_j}{\|y^r\|}.
\end{aligned}$$

As $\lambda_i^\infty>0$ for $i\in N_1$, $\lambda_j^\infty=0$ for $j\in N_0$ and all terms in the above expression are non-negative, we have

$$\sum_{i\in N_1}c_i/\|y^r\|\to 0 \tag{3.21}$$

$$\sum_{i\in N_1}\langle Q^i\cdot y^r,y^r\rangle/\|y^r\|\to 0 \tag{3.22}$$

$$\sum_{i\in N_0}\lambda_i^r\langle Q^i\cdot y^r,y^r\rangle/\|y^r\|\to 0 \tag{3.23}$$

<u>**Case1**</u> $N_1\neq\emptyset$.

We have $\|y^r\|\to\infty$ by (3.21) because $c_i>0$ for all $i\in N_1$ as 0 is a Slater point. By (3.22), it follows from the Schwart inequality,

$$\sum_{i\in N_1}\sqrt{\langle Q^i\cdot y^r,y^r\rangle}/\|y^r\|\leqslant |N_1|^{\frac{1}{2}}\left(\sum_{i\in N_1}\langle Q^i\cdot y^r,y^r\rangle/\|y^r\|^2\right)^{\frac{1}{2}}\to 0. \tag{3.24}$$

By dividing both sides of (3.20) by $\|y^r\|$, we have an absurdity that $1\leqslant 0$ after passing limits on both sides and making use of (3.21), (3.22), (3.23) and (3.24).

<u>**Case2**</u> $N_1=\emptyset$ i.e. $N_0=N$.

If $\{y^r\}_r$ is unbounded, then dividing both sides of (3.20) by $\|y^r\|$, we have an absurdity that $1\leqslant 0$ after passing limits on both sides and making use of (3.23). Therefore, $\{y^r\}_r$ must be bounded. However, $\lambda_i^r\to 0$ and $Q^i\cdot x^r=Q^i\cdot y^r$ for all $i\in N_0=N$, it follows that $\lambda_i^rQ^i\cdot x^r\to 0$ which contradicts (3.14). This completes the proof.

□

<u>**Proof of Theorem** 3.2.3</u>

For any $J \subseteq N \cup L = \{1, 2 \ldots, m\}$, let δ_J be the corresponding positive constant in Lemma 3.2.5 (where $\delta = \delta_J$, $N = N \cap J$, $L = L \cap J$). Define $\bar{\delta} := \min\{\delta_J \mid J \subseteq N \cup L\}$. Then $\bar{\delta} > 0$ since N is finite. Let $x \in S$, $I \subseteq \{1, 2 \ldots, m\}$, $\lambda_I > 0$ be such that $f_I(x) = 0$, $\|\sum_{i \in I} \lambda_i \nabla f_i(x)\| = 1$ and $\{\nabla f_i(x) \mid i \in I\}$ linearly independent. Then, by Lemma 3.2.5, we have

$$\frac{1}{\bar{\delta}} \geqslant \sum_{i \in N \cap I} \lambda_i + \left\| \sum_{j \in L \cap I} \lambda_j b^j \right\|.$$

Since $\{bj\}_{j \in L \cap I}$ is linearly independent and L is a finite set, it follows that $\|\lambda_I\|$ is uniformly bounded by a constant (independent on the choice of I and x). That is, f satisfies the (ACQ). Hence ,by Theorem 3.2.4, a global error bound exists.

□

3.3 Error Bounds for a Convex Inequality

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued closed proper convex function and assume that the closed convex set $S := f^{-1}(-\infty, 0]$ is non-empty throughout this section. By using convex analysis, we will present some necessary and sufficient conditions for the existence of a global error bound for the set S. That is, the existence of a positive constant r such that

$$dist(x, S) \leqslant r f(x)_+ \quad \text{for all } x \in \mathbb{R}^n. \tag{3.25}$$

After that, we will consider the set $S_C := C \cap f^{-1}(-\infty, 0]$, where C is a closed convex set in $\mathbb{R}^n$, and seek a global error bound of the following form:

$$dist(x, S_C) \leqslant r \max\{f(x)_+, dist(x, C)\} \quad \text{for all } x \in \mathbb{R}^n. \tag{3.26}$$

3.3.1 Unconstrained Case

Theorem 3.3.1 (Lewis and Pang [13]) *Let f and S be defined as above. For any positive constant $r > 0$, the following statements are equivalent:*

(a) The global error bound (3.25) *holds.*

(b) For all $\bar{x} \in f^{-1}(0)$ and $d \in N(\bar{x}; S)$, $f'(\bar{x}; d) \geqslant r^{-1}\|d\|$.

(c) For all $\bar{x} \in f^{-1}(0)$ and $d \in N(\bar{x}; S) \cap T(\bar{x}; dom(f))$, $f'(\bar{x}; d) \geqslant r^{-1}\|d\|$,

where N and T denote the normal cone and the tangent cone respectively.

Proof

(a) $\Rightarrow$ (b)

Let $\bar{x} \in f^{-1}(0)$ and $d \in N(\bar{x}; S)$. Then, if $x \in S$, one has that $\langle d, x - \bar{x}\rangle \leqslant 0$. In view of the variational inequality, it follows that $dist(\bar{x} + \tau d, S) = \|\tau d\|$ for all $\tau > 0$ because $\langle \bar{x} + \tau d - \bar{x}, x - \bar{x}\rangle \leqslant 0$ for any $x \in S$ and $\tau > 0$. By (a), it follows that

$$r(f(\bar{x} + \tau d) - f(\bar{x})) = rf(\bar{x} + \tau d) \geqslant dist(\bar{x} + \tau d, S) = \tau\|d\|.$$

Dividing both sides by τ and let $\tau \to 0^+$, we get $rf'(\bar{x}; d) \geqslant \|d\|$ and so (b) holds.

(b) $\Rightarrow$ (c)

This is obvious.

(c) $\Rightarrow$ (a)

Let $x \in \mathbb{R}^n$. To prove (3.25), we may suppose that $x \in dom(f) \backslash S$. Let $\bar{x}$ be the projection of x into S. Then, by the variational inequality, $x - \bar{x} \in N(\bar{x}; S)$. Since $dom(f)$ is closed and convex, $x - \bar{x} \in \overline{cone(dom(f) - \bar{x})} = T(\bar{x}; dom(f))$ and hence $x - \bar{x} \in N(\bar{x}; S) \cap T(\bar{x}; dom(f))$. We claim that $f(\bar{x}) = 0$. Suppose on the contrary that $f(\bar{x}) < 0$. Then $f(\bar{x} + \tau(x - \bar{x})) \leqslant \tau f(x) + (1 - \tau) f(\bar{x}) < 0$ for $\tau > 0$ and small enough. It follows that $\bar{x} + \tau(x - \bar{x}) \in S$ with

$$dist(x, \bar{x} + \tau(x - \bar{x})) = (1 - \tau)\|x - \bar{x}\| < \|x - \bar{x}\| = dist(x, S),$$

which is impossible. Therefore, $\bar{x} \in f^{-1}(0)$ and it follows from 3 and the convexity of f that

$$\begin{aligned} f(x)_+ = f(x) &= f(x) - f(\bar{x}) \\ &\geqslant f'(\bar{x}; x - \bar{x}) \\ &\geqslant r^{-1}\|x - \bar{x}\| \\ &= r^{-1} dist(x, S), \end{aligned}$$

verifying (a).

□

Before going to prove some sufficient conditions for (3.25) holds, we need the following lemma.

Lemma 3.3.1 *Let f and S be defined as in the beginning. Let L be the linear subspace of $\mathbb{R}^n$ parallel to the affine hull of dom(f). If the Slater condition holds (That is, $f(x^\star) < 0$ for some $x^\star \in \mathbb{R}^n$.), then for every $\bar{x} \in f^{-1}(0) \cap ri(dom(f))$,*

$$N(\bar{x}; S) \cap L = cone(\partial f(\bar{x})) \cap L;$$

consequently,

$$N(\bar{x}; S) \cap T(\bar{x}; dom(f)) \subseteq cone(\partial f(\bar{x})).$$

Proof

The second conclusion follows from the first because $T(\bar{x}; dom(f)) \subseteq L$ since $\bar{x} - dom(f) \in L$. Without loss of generality, we assume $\bar{x} = 0$. Since $0 \in ri(dom(f)) \subseteq dom(f)$, we have $L =$aff$(dom(f))$ and write $\mathbb{R}^n = L \oplus L^\perp$, where $L^\perp$ is the orthogonal complement of L. Define $g := f_L$, the restriction of f on L. Then g is a closed proper convex function on L satisfying the Slater condition. Moreover, the origin is an interior point of $dom(g)$ in the space L since $0 \in ri(dom(f)) = int(dom(g))$ in the space L. By Corollary 2.2.1, we have

$L \cap N(0; g^{-1}(-\infty, 0]) = L \cap cone(\partial g(0))$ and hence,

$$L \cap N(0; g^{-1}(-\infty, 0]) \subseteq cone(\partial f(0)) \cap L. \tag{3.27}$$

We also claim that

$$N(0; S) = \left(L \cap N(0; g^{-1}(-\infty, 0])\right) \oplus L^{\perp}. \tag{3.28}$$

To show this, let $\hat{x} \in N(0; S)$ and write $\hat{x} = x_1 + x_2$, where $x_1 \in L$ and $x_2 \in L^{\perp}$. Then

$$\langle x_1 + x_2, x \rangle \leqslant 0 \quad \text{for all } x \in S = f^{-1}(-\infty, 0] = g^{-1}(-\infty, 0] \subseteq L;$$

hence

$$\langle x_1, x \rangle \leqslant 0 \quad \text{for all } x \in g^{-1}(-\infty, 0].$$

Therefore, $x_1 \in N\left(0; g^{-1}(-\infty, 0]\right)$ and hence $\hat{x} \in (L \cap N(0; g^{-1}(-\infty, 0])) \oplus L^{\perp}$. Similarly, we can prove the reverse inclusion.
By (3.27) and (3.28), we have

$$\begin{aligned} N(0; S) \cap L &= \left((L \cap N(0; g^{-1}(-\infty, 0])) \oplus L^{\perp}\right) \cap L \\ &\subseteq N(0; g^{-1}(-\infty, 0]) \cap L \\ &\subseteq cone(\partial f(0)) \cap L, \end{aligned}$$

and hence the equalities hold throughout because $\partial f(0) \subseteq N(0; S)$ by definition.
□

Proposition 3.3.1 *Let f and S be defined as in the beginning, the Slater condition holds if and only if $0 \notin \partial f(f^{-1}(0))$.*

Proof
Suppose $0 \in \partial f(f^{-1}(0))$: $f(y) - f(\bar{x}) \geqslant \langle 0, y - \bar{x} \rangle$ for some $\bar{x} \in f^{-1}(0)$, for all

$y \in \mathbb{R}^n$. Therefore, $f(y) \geqslant 0$ for all $y \in \mathbb{R}^n$ and hence the Slater condition fails. Conversely, suppose the Slate condition does hold: $f(y) \geqslant 0$ for all $y \in \mathbb{R}^n$. Then, since S is assumed to be non-empty, there exists x_0 such that $f(x_0) = 0$. Clearly, $0 \in \partial f(x_0) \subseteq \partial f(f^{-1}(0))$.

□

Corollary 3.3.1 *Let f and S be defined as in the beginning. Assume that $f^{-1}(0) \subseteq ri(dom(f))$. Consider the following statements:*

(a) (Global error bound) the global error (3.25) holds for some $r > 0$;

(b) (Strong Slater) $0 \notin \overline{\partial f(f^{-1}(0))}$;

(c) (Well-posedness and Slater) f is well-posed and satisfies the Slater condition;

(d) (Slater and a subcase of well-posedness) $0 \in ri(dom(f^\star)) \backslash \partial f(f^{-1}(0))$;

(e) (Recession) $0 \notin \overline{dom(f^\star)}$, or equivalently, there exists an $d \in \mathbb{R}^n$ such that $f_\infty(d) < 0$;

(f) (Slater and positive homogeneity) f is positively homogeneous and satisfies the Slater condition.

Then the following implications hold:

$$\begin{array}{ccccccc} (f) & \Rightarrow & (e) & \Rightarrow & (b) & \Rightarrow & (a) \\ & & & & \Uparrow & & \\ & & & & (d) \Rightarrow (c) & & \end{array}$$

__Proof__

__(b)⇒(a)__

If (b) holds, then so does the Slater condition by Proposition 3.3.1. Let $\delta := dist(0, \overline{\partial f(f^{-1}(0))}) > 0$. For any $\bar{x} \in f^{-1}(0) = f^{-1}(0) \cap ri(dom(f))$ and $d \in$

$N(\bar{x}; S) \cap T(\bar{x}; dom(f))$, we have $d \in cone(\partial f(\bar{x}))$ by Lemma 3.3.1. Hence $\frac{d}{\lambda} \in \partial f(\bar{x})$ for some $\lambda > 0$; note then that $\|\frac{d}{\lambda}\| \geqslant \delta$ by the definition of δ. Consequently, it follows from Theorem 2.2.2 that

$$f^{'}(\bar{x}; d) \geqslant \langle d/\lambda, d \rangle = \frac{1}{\lambda}\|d\|^2 \geqslant \delta\|d\|.$$

By Theorem 3.3.1, (a) holds with $r = \delta^{-1}$.

(c)⇒(b)

Suppose (b) fails: There exists sequences $\{ \bar{x}^k \} \subset f^{-1}(0)$ and $\{ d^k \}$ with $d^k \in \partial f(\bar{x}^k)$ for each k and $d^k \to 0$. Thus $\{ \bar{x}^k \}$ is a stationary sequence and hence is minimizing by the well-posedness of f. So $\inf_{x \in \mathbb{R}^n} f(x) = \lim_{x \to \infty} f(\bar{x}^k) = 0$. This contradicts the Slater condition.

(d)⇒(c)

It follows from Theorem 2.3.4 and Proposition 3.3.1 immediately.

(e)⇒(b)

The equivalence stated in (e) follows from Corollary 2.1.1. To show (e) ⇒ (b), it suffices to prove that $\partial f(f^{-1}(0)) \subseteq dom(f^\star)$. To see the inclusion, let $u \in \partial f(x_0)$ for some $x_0 \in f^{-1}(0)$. Then

$$\langle u, y \rangle - \langle u, x_0 \rangle \leqslant f(y) - f(x_0) = f(y), \quad \text{for all } y \in \mathbb{R}^n.$$

It implies that

$$f^\star(u) = \sup\{ \langle u, y \rangle - f(y) \mid y \in \mathbb{R}^n \} \leqslant \langle u, x_0 \rangle < \infty.$$

Hence, $u \in dom(f^\star)$.

(f)⇒(e)

By the Slater condition, let $x^\star \in \mathbb{R}^n$ satisfy $f(x^\star) < 0$. Since f is positively homogeneous, it follows that

$$f_\infty(x^\star) = \lim_{t \to \infty} \frac{f(0 + tx^\star) - f(0)}{t} = \lim_{t \to \infty} \frac{tf(x^\star)}{t} = f(x^\star) < 0.$$

Hence, (e) holds.

□

Remark: The condition '$f^{-1}(0) \subseteq ri(dom(f))$' is automatically satisfied when f is finite-valued. In this case, $ri(dom(f)) = \mathbb{R}^n$.

3.3.2 Constrained Case

Theorem 3.3.2 (Lewis and Pang [13]) *Let f, C and S_C be defined as in the beginning. Moreover, suppose C satisfies either,*

(a) the projection condition: $\Pi_C(dom(f)) \subseteq dom(f)$, or

(b) the interiority condition: $\{\, x \in C \mid f(x) < 0 \,\} \subseteq int(dom(f))$.

For any positive constant $r \geqslant 1$, the global error bound (3.26) holds if and only if for any $\bar{x} \in f^{-1}(0) \cap C$ and $d \in N(\bar{x}; S_C)$,

$$\max\{\, dist(d, T(\bar{x}, C)), f'(\bar{x}; d) \,\} \geqslant r^{-1}\|d\| \tag{3.29}$$

Proof

Let $g(x) = dist(x, C) = \|x - \Pi_C(x)\|$ for all $x \in \mathbb{R}^n$. Then g is convex as C is convex. For any $\bar{x} \in C$ and $d \in \mathbb{R}^n$,

$$\begin{aligned} g'(\bar{x}; d) &= \lim_{t \downarrow 0} \frac{g(\bar{x} + td) - g(\bar{x})}{t} \\ &= \lim_{t \downarrow 0} \frac{\|\bar{x} + td - \Pi_C(\bar{x} + td)\|}{t} \\ &= \lim_{t \downarrow 0} \left\| d - \frac{\Pi_C(\bar{x} + td) - \Pi_C(\bar{x})}{t} \right\| \\ &= \|d - \Pi'_C(\bar{x}; d)\| \\ &= dist(d, T(\bar{x}; C)), \end{aligned}$$

where the last equality follows from the fact that $\Pi'_C(\bar{x}; d) = \Pi_{T(\bar{x};C)}(d)$ by Theorem 2.2.8. Consider the convex function given by

$$\bar{f}(x) := \max\{\, f(x), g(x) \,\} \quad \text{for } x \in \mathbb{R}^n.$$

It is easy to see that $S_C = \bar{f}^{-1}(-\infty, 0]$ and $\bar{f}(x)_+ = \max\{\, f(x)_+, dist(x, C)\,\}$. Moreover, for any $\bar{x} \in f^{-1}(0) \cap C$ and $d \in \mathbb{R}^n$,

$$\begin{aligned}
\bar{f}'(\bar{x}; d) &= \lim_{t \downarrow 0} \frac{\bar{f}(\bar{x} + td) - \bar{f}(\bar{x})}{t} \\
&= \lim_{t \downarrow 0} \frac{\bar{f}(\bar{x} + td)}{t} \\
&= \lim_{t \downarrow 0} \frac{f(\bar{x} + td) + g(\bar{x} + td) + |f(\bar{x} + td) - g(\bar{x} + td)|}{2t} \\
&= \lim_{t \downarrow 0} \Big(\frac{f(\bar{x} + td) - f(\bar{x}) + g(\bar{x} + td) - g(\bar{x})}{2t} + \\
&\qquad \frac{|f(\bar{x} + td) - f(\bar{x}) - g(\bar{x} + td) + g(\bar{x})|}{2t} \Big) \\
&= \frac{f'(\bar{x}; d) + g'(\bar{x}; d) + \left| f'(\bar{x}; d) - g'(\bar{x}; d) \right|}{2} \\
&= \max\{\, f'(\bar{x}; d), g'(\bar{x}; d)\,\} \\
&= \max\{\, f'(\bar{x}; d), dist(d, T(\bar{x}; C))\,\}.
\end{aligned}$$

Suppose the error bound (3.26) holds: $dist(x, S_C) \leqslant r\bar{f}(x)_+$. By Theorem 3.3.1, $\bar{f}'(\bar{x}; d) \geqslant r^{-1}\|d\|$ for any $\bar{x} \in \bar{f}^{-1}(0)$ and $d \in N(\bar{x}; S_C)$. Since $f^{-1}(0) \cap C \subseteq \bar{f}^{-1}(0)$, this proves the necessity part of the theorem.

Conversely, suppose (3.29) holds for some $r \geqslant 1$ and all $(\bar{x}, d) \in (f^{-1}(0) \cap C) \times N(\bar{x}; S_C)$. For any $x \in dom(f)$ and let $\bar{x} = \Pi_{S_C}(x)$, we have two cases:

<u>Case 1</u> $f(\bar{x}) = 0$.

In this case, $\bar{f}(\bar{x}) = \max\{\, f(\bar{x}), dist(\bar{x}, C)\,\} = 0$ as $\bar{x} \in C$. That is, $\bar{x} \in \bar{f}^{-1}(0)$. Following the third part of proof of the Theorem 3.3.1, we have

$$dist(x, S_C) \leqslant r\bar{f}(x)_+ = r\max\{\, f(x)_+, dist(x, C)\,\},$$

where $f = \bar{f}$, $S = S_C$.

<u>Case 2</u> $f(\bar{x}) < 0$.

We claim that $\Pi_C(x) = \bar{x}$. Suppose not, let $x' = \Pi_C(x) \neq \bar{x}$. Then $\|x - x'\| < \|x - \bar{x}\|$. If the projection condition holds, then $x' \in dom(f)$. If the interiority condition holds, then $\bar{x} \in int(dom(f))$. For both cases, $f(\bar{x} + \tau(x' - \bar{x})) < 0$ for

$\tau > 0$ and small enough. Hence, $\bar{x} + \tau(x' - \bar{x}) \in S_C$.

$$\begin{aligned}\|x - (\bar{x} + \tau(x' - \bar{x})\| &\leqslant \tau\|x - x'\| + (1-\tau)\|x - \bar{x}\| \\ &< \tau\|x - \bar{x}\| + (1-\tau)\|x - \bar{x}\| \\ &= \|x - \bar{x}\|,\end{aligned}$$

contradicting to the definition of $\bar{x}$. So $\Pi_C(x) = \bar{x}$ and hence

$$dist(x, S_C) = dist(x, C) \leqslant r \max\{ f(x)_+, dist(x, C) \},$$

where the last inequality holds since $r \geqslant 1$ by assumption.

□

Remark: The projection conditions is trivial if $C = \mathbb{R}^n$ or f is finite-valued. Note that for any $\bar{x} \in C \cap f^{-1}(0)$,

$$N(\bar{x}; S_C) \supseteq N(\bar{x}; C) + N(\bar{x}; f^{-1}(-\infty, 0]) \supseteq N(\bar{x}; C) + cone(\partial f(\bar{x})).$$

Under the Slater condition and $C \cap f^{-1}(0) \subseteq int(dom(f))$, the following lemma shows that equality holds throughout the above expression.

Lemma 3.3.2 *Suppose that*

(a) $C \cap f^{-1}(0) \subseteq int(dom(f))$;

(b) the Slater condition holds for S_C;

(c) the interiority condition holds.

Then, for any $\bar{x} \in C \cap f^{-1}(0)$,

$$N(\bar{x}; S_C) = N(\bar{x}; C) + N(\bar{x}; f^{-1}(-\infty, 0]) = N(\bar{x}; C) + cone(\partial f(\bar{x})).$$

Proof

Define $f_1 := I_C$, $f_2 := I_S$ and $f_0 := f_1 + f_2$, we have

$$dom(f_1) = C,\ dom(f_2) = S,\ dom(f_0) = S_C \text{ and } f_0 = I_{S_C}.$$

Let $\hat{x} \in S_C$ be a Slater point. By the interiority condition, f is finite on some neighborhood of $\hat{x}$ and hence is strictly smaller than zero on some neighborhood $\mathbb{E} \subseteq int(dom(f_2))$ of $\hat{x}$ as f is convex and $f(\hat{x}) < 0$. Therefore,

$$\hat{x} \in dom(f_1) \cap int(dom(f_2)) \neq \emptyset.$$

Clearly, $\bar{x} \in dom(f_1) \cap dom(f_2)$. By Theorem 2.2.14, we have

$$\partial f_0(\bar{x}) = \partial f_1(\bar{x}) + \partial f_2(\bar{x}).$$

Finally, by Theorem 2.2.9, the first equality follows.

Since $\bar{x} \in C \cap f^{-1}(0) \subseteq int(dom(f))$ and the Slater condition holds, by Corollary 2.2.1, we have $N(\bar{x}; f^{-1}(-\infty, 0]) = cone(\partial f(\bar{x}))$. Hence, the second equality follows.

□

Proposition 3.3.2 *If $\bar{x} \in \mathbb{R}^n$ is such that $N(\bar{x}; S_C) = N(\bar{x}; C) + cone(\partial f(\bar{x}))$, then for all non-zero vector $d \in N(\bar{x}; S_C)$, we have*

$$\max\{\, dist(d, T(\bar{x}; C)), f^{'}(\bar{x}; d) \,\} > 0.$$

Proof

Suppose the contrary that $\max\{\, dist(d, T(\bar{x}; C)), f^{'}(\bar{x}; d) \,\} = 0$ for some $\bar{x} \in \mathbb{R}^n$ with the above property and some $d \in N(\bar{x}; S_C) \backslash \{\, 0 \,\}$. Then $d \in T(\bar{x}; C)$. Write $d = u + \lambda v$, where $(u, v) \in N(\bar{x}; C) \times \partial f(\bar{x})$ and $\lambda \geqslant 0$. By Theorem 2.2.2, we have a contradiction that,

$$0 < \langle d, d \rangle = \langle d, u \rangle + \lambda \langle d, v \rangle \leqslant \lambda \langle d, v \rangle \leqslant \lambda f^{'}(\bar{x}; d) \leqslant 0,$$

where the second inequality is due to the fact that $T(\bar{x}; C)$ is the polar of $N(\bar{x}; C)$.

□

Corollary 3.3.2 *Let f, C and S_C be defined as in the beginning. Moreover, suppose C satisfies the interiority condition and $C \cap f^{-1}(0) \subseteq int(dom(f))$. Consider the following statements:*

(a) (Global error bound) there exists a constant $r > 0$ such that the global error bound (3.26) holds;

(b) (Strong Slater) there exists a constant $K > 0$ such that for any $\bar{x} \in C \cap f^{-1}(0)$, $\lambda > 0$ and $(u, v) \in N(\bar{x}; C) \times \partial f(\bar{x})$, one has $\|u\| + \lambda \leqslant K\|u + \lambda v\|$;

(c) (Bounded subgradients and a constraint qualification) the set $\partial f(C \cap f^{-1}(0))$ is bounded and $\inf\{\, dist(-N(\bar{x}; C), \partial f(\bar{x})) \mid \bar{x} \in C \cap f^{-1}(0)\,\} > 0$;

(d) (Generalized Robinson condition) the Slater condition holds and $\overline{C \cap f^{-1}(0)}$ is bounded and contained in $int(dom(f))$.

The following implications hold:

$$(d) \Rightarrow (c) \Rightarrow (b) \Rightarrow (a).$$

Proof

(b) $\Rightarrow$ (a)

Suppose on the contrary (by Theorem 3.3.2) that there exist sequences $\{\, x^k \,\} \subseteq C \cap f^{-1}(0)$ and $\{\, d^k \,\} \subseteq N(x^k; S_C) \backslash \{\, 0 \,\}$ such that for each k,

$$\max\{\, dist(d^k, T(x^k; C)), f'(x^k; d^k) \,\} \leqslant \|d^k\|/k. \tag{3.30}$$

By homogeneity, we may assume that each d^k is a unit vector. For each k, let $w^k = \Pi_{T(x^k;C)}(d^k)$. Then

$$\lim_{k \to \infty} \|d^k - w^k\| = 0. \tag{3.31}$$

Note that under Assumption (b), the Slater condition holds for S_C (see remark below). Hence, $N(x^k; S_C) = N(x^k; C) + cone(\partial f(x^k))$ by Lemma 3.3.2. Therefore, we may write $d^k = u^k + \lambda_k v^k$ for some $(u^k, v^k) \in N(x^k; C) \times \partial f(x^k)$ and $\lambda_k \geqslant 0$. By Assumption (b), $\|u^k\| + \lambda_k \leqslant K$. Therefore, $\{\, \lambda_k \,\}$ and $\{\, u^k \,\}$ are bounded. On the other hand, $\langle d^k, v^k \rangle \leqslant f'(x^k; d^k) \leqslant 1/k$ by (3.30), so

$$\limsup_{k \to \infty} \lambda_k \langle d^k, v^k \rangle \leqslant 0. \tag{3.32}$$

We have

$$\begin{aligned}1 = \langle d^k, d^k\rangle &= \langle d^k, u^k\rangle + \lambda_k\langle d^k, v^k\rangle \\ &= \langle d^k - w^k, u^k\rangle + \langle w^k, u^k\rangle + \lambda_k\langle d^k, v^k\rangle \\ &\leqslant \langle d^k - w^k, u^k\rangle + \lambda_k\langle d^k, v^k\rangle,\end{aligned}$$

since w^k, u^k are respectively from the tangent and the normal cones of C at x^k. This leads to a contradiction by taking lim sup on both sides since the lim sup of the right-hand side is non-positive by (3.31) and (3.32).

<u>**(c) ⇒ (b)**</u>

Assume (b) fails. Then take sequences $\{x^k\} \subseteq C \cap f^{-1}(0)$, $\{(u^k, v^k)\} \subseteq N(x^k; C) \times \partial f(x^k)$ and $\{\lambda_k\} \subset \mathbb{R}$ with $\lambda_k > 0$ for each k such that

$$\|u^k + \lambda_k v^k\| < k^{-1}(\|u^k\| + \lambda_k).$$

By Assumption (c), $\{v^k\}$ is bounded, hence so is the sequence $\{\lambda_k^{-1}u^k\}$ by the above inequality. Let $\bar{u}^k = \lambda_k^{-1}u^k$. Then $\bar{u}^k \in N(x^k; C)$ for each k and

$$\|\bar{u}^k + v^k\| < k^{-1}(\|\bar{u}^k\| + 1).$$

Let $k \to \infty$, we have $\lim_{k\to\infty} \|\bar{u}^k + v^k\| = 0$, which contradicts to the assumption that $\inf\{ dist(-N(\bar{x}; C), \partial f(\bar{x})) \mid \bar{x} \in C \cap f^{-1}(0)\} > 0$.

<u>**(d) ⇒ (c)**</u>

Under Assumption (d), the boundness of $\partial f(C \cap f^{-1}(0))$ follows from Theorem 2.2.17. To complete the proof, suppose the contrary that there exists sequences $\{x^k\} \subseteq C \cap f^{-1}(0)$ and $\{(u^k, v^k)\}$ with $(u^k, v^k) \in N(x^k; C) \times \partial f(x^k)$ for each k such that $\|u^k + v^k\| \to 0$ as $k \to \infty$. Let $x^\star \in C$ be a Slater point: $f(x^\star) < 0$. We have $\langle x^\star - x^k, u^k + v^k\rangle \leqslant f(x^\star) - f(x^k) = f(x^\star)$. Since $\{x^k\}$ is bounded by assumption, the left-hand side of the above inequality converges to 0, so $0 \leqslant f(x^\star)$ which contradicts to $f(x^\star) < 0$. Consequently, (c) holds under Assumption (d).

□

Remark: Under Assumption (b), the Slater condition must hold for the set S_C.

Proof of the Remark

Suppose on the contrary that the Slater condition fails. Then

$$S_C = C \cap f^{-1}(-\infty, 0] = C \cap f^{-1}(0).$$

It follows that any $x \in S_C$ is an global minimizer of f on C. Hence, for any $x \in S_C$, we have $0 \in \partial(f(x) + I_C(x))$. Note that $x \in dom(I_C) \cap dom(f)$ and

$$dom(I_C) \cap int(dom(f)) = C \cap int(dom(f)) \neq \emptyset$$

as $C \cap f^{-1}(0) \subseteq int(dom(f))$. By Theorem 2.2.9 and Theorem 2.2.14, we have

$$0 \in \partial(f(x) + I_C(x)) = \partial f(x) + \partial I_C(x) = \partial f(x) + N(x; C).$$

Write $0 = u + v$ where $(u, v) \in N(x; C) \times \partial f(x)$. But this contradicts (b).

□

Definition 3.3.1 *S_C is said to be metrically regular at a vector $\bar{x} \in S_C$ if there exist positive constants r and δ such that*

$$dist(y, S_C) \leqslant r \max\{\, f(y)_+, dist(y, C)\,\} \quad \text{for all } y \in B(\bar{x}, \delta).$$

That is, a local error bound exists.

Now, we are going to prove a necessary condition for the existence of a local error bound for the set S_C.

Proposition 3.3.3 (Lewis and Pang [13]) *Let $\bar{x} \in C \cap f^{-1}(0)$ and assume $\bar{x} \in dom(\partial f)$. If S_C is metrically regular at $\bar{x}$, then*

$$N(\bar{x}; S_C) = \overline{N(\bar{x}; C) + cone(\partial f(\bar{x}))}.$$

If in addition $\bar{x} \in int(dom(f))$, then the closure can be dropped from the right-hand side.

Proof

$N(\bar{x}; S_C) \supseteq \overline{N(\bar{x}; C) + cone(\partial f(\bar{x}))}$ obviously. The proof for the reverse inclusion is contained in 2 cases.

Case 1 $C = \mathbb{R}^n$, we have $S_C = S$.

It suffices to show that $N(\bar{x}; S) \subseteq \overline{cone(\partial f(\bar{x}))}$. By Theorem 2.2.12, $N(\bar{x}; S) = \overline{cone(\partial f(\bar{x}))}$ if the Slater condition holds. Moreover, if $\bar{x} \in int(dom(f))$, the closure can be dropped by Corollary 2.2.1. Therefore, it remains to consider the case when the Slater condition fails to satisfy: $f(x) \geqslant 0$ for all $x \in dom(f)$. For such a case, let r be the positive constant in Definition 3.3.1. Then, by the metrically regularity, we have

$$dist(y, S) \leqslant rf(y) \quad \text{for all } y \text{ sufficiently close to } \bar{x}. \tag{3.33}$$

Fix any $\bar{y} \in N(\bar{x}; S)$ and note that $\bar{x}$ is a global minimizer of the function $\langle -\bar{y}, x\rangle$ on S. By Theorem 2.4.1, there exists a positive constant α such that $\bar{x}$ is an unconstrained global minimizer of the following function:

$$x \mapsto \langle -\bar{y}, x\rangle + \alpha dist(x, S).$$

Therefore, for any x_0 sufficiently closed to $\bar{x}$, we have by (3.33) that

$$\begin{aligned}\langle -\bar{y}, \bar{x}\rangle &= \langle -\bar{y}, \bar{x}\rangle + \alpha dist(\bar{x}, S)\\ &\leqslant \langle -\bar{y}, x_0\rangle + \alpha dist(x_0, S)\\ &\leqslant \langle -\bar{y}, x_0\rangle + \alpha r f(x_0).\end{aligned}$$

That is, $(\alpha r)^{-1}\langle \bar{y}, x_0 - \bar{x}\rangle \leqslant f(x_0)$ for x_0 sufficiently closed to $\bar{x}$. For any $x \in \mathbb{R}^n$, letting $x_0 = (1-\tau)\bar{x} + \tau x$ with $\tau > 0$ sufficiently small and making use of the convexity of f, it follows that

$$(\alpha r)^{-1}\langle \bar{y}, \tau(x - \bar{x})\rangle \leqslant f((1-\tau)\bar{x} + \tau x) \leqslant \tau f(x),$$

showing that $(\alpha r)^{-1}\bar{y} \in \partial f(\bar{x})$, hence that $N(\bar{x}; S) \subseteq cone(\partial f(\bar{x}))$.

Case 2 C is a proper subset of $\mathbb{R}^n$.

Define $\bar{f}(x) := \max\{\, f(x), dist(x, C)\,\}$ for all $x \in \mathbb{R}^n$.

By the similar argument above,

$$N(\bar{x}; S_C) = \overline{cone(\partial \bar{f}(\bar{x}))}.$$

By Theorem 2.2.10 and Theorem 2.2.15,

$$\partial \bar{f}(\bar{x}) = co(\partial f(\bar{x}) \cup (N(\bar{x}; C) \cap \mathbb{B})),$$

where $\mathbb{B}$ is the closed unit ball in $\mathbb{R}^n$. Thus,

$$cone(\partial \bar{f}(\bar{x})) = N(\bar{x}; C) + cone(\partial f(\bar{x}))$$

and so

$$N(\bar{x}; S_C) = \overline{N(\bar{x}; C) + cone(\partial f(\bar{x}))}.$$

Note that the closure can be dropped if $\bar{x} \in int(dom(f))$ as $dom(f) = dom(\bar{f})$. This completes the proof.

□

3.4 Error Bounds for System of Convex Inequalities

In this section, we study some sufficient conditions for the existence of an error bound for a system of convex inequalities. Let X be a real reflexive Banach space and C be a non-empty closed convex subset of X. Consider the convex set

$$\begin{aligned} S_C :&= \{\, x \in C \mid f_i(x) \leqslant 0 \text{ for all } i = 1, \ldots, m \,\} \quad \text{or simply} \\ &= \{\, x \in C \mid F(x) \leqslant 0 \,\}, \end{aligned}$$

where $F(x) = (f_1(x), \ldots, f_m(x))$ is a vector-valued function from X to $\mathbb{R}^m$ with each f_i is a real-valued convex function on X. We always assume that S_C is non-empty throughout this section.

3.4.1 Unconstrained Case

In this subsection, we only consider $C = X = \mathbb{R}^n$, $\|\cdot\|$ to be an arbitrary norm on $\mathbb{R}^n$ and present an global error bound of the following form under a Strong Slater constraint qualification:

$$dist(x, S) \leqslant \hat{r}\|F(x)_+\|_\infty \quad \text{for all } x \in \mathbb{R}^n, \tag{3.34}$$

where $\hat{r}$ is a positive number and the distance function is defined with respect to the norm $\|\cdot\|$.

Definition 3.4.1 (Slater constraint qualification (CQ)) *F is said to satisfy the Slater CQ if there exists an $\hat{x} \in \mathbb{R}^n$ such that*

$$f_i(\hat{x}) < 0 \quad \textit{for } i = 1, \ldots, m.$$

We call $\hat{x}$ a Slater point of F and use S^0 to denote the set of all such points.

Definition 3.4.2 (Strong Slater CQ) *F is said to satisfy the Strong Slater CQ if*

(a) F satisfies the Slater CQ and

(b) there exists an $r > 0$ such that

$$\sup_{p \in \partial S} \inf_{\hat{x} \in S^0} \frac{\|\hat{x} - p\|}{\min_{1 \leqslant i \leqslant m} -f_i(\hat{x})} \leqslant r < \infty,$$

where $\partial S := S \backslash S^0$.

Lemma 3.4.1 *For $\hat{r} > r$, the Strong Slater CQ implies that*

$$\textit{for all } p \in \partial S, \textit{ there exists an } \hat{x}(p) \in S^0 \textit{ such that}$$
$$\frac{\|\hat{x}(p) - p\|}{\min_{1 \leqslant i \leqslant m} -f_i(\hat{x}(p))} \leqslant \hat{r} < \infty \tag{3.35}$$

Conversely, (3.35) implies the Strong Slater CQ with $r = \hat{r}$.

Proof

Suppose the Strong Slater CQ holds. Then

$$\inf_{\hat{x}\in S^0} \frac{\|\hat{x}-p\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x})} \leqslant r < \infty \quad \text{for all } p \in \partial S.$$

For any $\hat{r} > r$, it follows from the property of infimum that there exists an $\hat{x}(p) \in S^0$ satisfying

$$\begin{aligned}
\frac{\|\hat{x}(p)-p\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x}(p))} &< \inf_{\hat{x}\in S^0} \frac{\|\hat{x}-p\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x})} + (\hat{r}-r) \\
&\leqslant \inf_{\hat{x}\in S^0} \frac{\|\hat{x}-p\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x})} + (\hat{r} - \inf_{\hat{x}\in S^0} \frac{\|\hat{x}-p\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x})}) \\
&= \hat{r}.
\end{aligned}$$

Conversely, suppose that (3.35) holds. Then for all $p \in \partial S$, we have

$$\inf_{\hat{x}\in S^0} \frac{\|\hat{x}-p\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x})} \leqslant \hat{r} < \infty.$$

Thus

$$\sup_{p\in\partial S} \inf_{\hat{x}\in S^0} \frac{\|\hat{x}-p\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x})} \leqslant \hat{r} < \infty.$$

□

Theorem 3.4.1 (Mangasarian [16]) *Under the Strong Slater CQ, the global error bound (3.34) holds for any $\hat{r} > r$.*

Before proving the above theorem, we need a lemma and some theorems first. Now, let us consider the following constrained minimization problem (CMP):

$$\min_{x\in S} \theta(x),$$

where θ is a real-valued convex function on $\mathbb{R}^n$.

Theorem 3.4.2 *[17, Thm. 5.4.7] Suppose $\bar{x} \in \mathbb{R}^n$ is a solution of the (CMP). Then there exists an $\bar{u} = (\bar{u_1}, \ldots, \bar{u_m}) \in \mathbb{R}^m$ with each $\bar{u_i} \geqslant 0$ such that* *(a)* $\langle \bar{u}, F(\bar{x})\rangle = 0$ *and*

(b) $\theta(\bar{x}) + \langle u, F(\bar{x})\rangle \leqslant \theta(\bar{x}) \leqslant \theta(x) + \langle \bar{u}, F(x)\rangle$ *for all* $u \geqslant 0$ *in* $\mathbb{R}^m$ *and* $x \in \mathbb{R}^n$.

Lemma 3.4.2 *Let* $\hat{x} \in S^0$*. Then for any* $x \in \mathbb{R}^n$*, there exists an* $\bar{u} \geqslant 0$ *in* $\mathbb{R}^m$ *such that*

$$\|\bar{u}\|_1 \leqslant \frac{\|x - \hat{x}\| - \|x - \Pi_S(x)\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x})}. \tag{3.36}$$

Proof

By Theorem 3.4.2, there exists an $\bar{u} \geqslant 0$ in $\mathbb{R}^m$ such that

$$\|\Pi_S(x) - x\| \leqslant \|\hat{x} - x\| + \langle \bar{u}, F(\hat{x})\rangle.$$

Hence,

$$\|\bar{u}\|_1 \cdot \min_{1\leqslant i\leqslant m} -f_i(\hat{x}) \leqslant -\langle \bar{u}, F(\hat{x})\rangle \leqslant \|\hat{x} - x\| - \|\Pi_S(x) - x\|.$$

Dividing both sides by $\min_{1\leqslant i\leqslant m} -f_i(\hat{x})$, the proof is completed.

□

Theorem 3.4.3 *Let* $\hat{x} \in S^0$*. For any* $x \in \mathbb{R}^n$ *with* $x \neq \hat{x}$*, we have*

$$dist(x, S) \leqslant \frac{\|x - \hat{x}\| \cdot \|F(x)_+\|_\infty}{\|F(x)_+\|_\infty + \min_{1\leqslant i\leqslant m} -f_i(\hat{x})}.$$

Proof

By Theorem 3.4.2, there exists an $\bar{u} \geqslant 0$ in $\mathbb{R}^m$ such that

$$dist(x, S) = \|x - \Pi_S(x)\| \leqslant \|x - x\| + \langle \bar{u}, F(x)\rangle \leqslant \|\bar{u}\|_1 \cdot \|F(x)_+\|_\infty.$$

It follows from Lemma 3.4.2 that

$$\|x - \Pi_S(x)\| \leqslant \frac{\|x - \hat{x}\| - \|x - \Pi_S(x)\|}{\min_{1\leqslant i\leqslant m} -f_i(\hat{x})}\|F(x)_+\|_\infty \tag{3.37}$$

and hence

$$\|x - \Pi_S(x)\| \leqslant \frac{\|x - \hat{x}\| \cdot \|F(x)_+\|_\infty}{\|F(x)_+\|_\infty + \min_{1 \leqslant i \leqslant m} -f_i(\hat{x})}.$$

□

Proof of Theorem 3.4.1

Without loss of generality, we can assume that $x \notin S$. Note that $\Pi_S(x) \in \partial S$. It follows from Lemma 3.4.1 that there exists an $\hat{x}(\Pi_S(x)) \in S^0$ such that

$$\frac{\|\hat{x}(\Pi_S(x)) - \Pi_S(x)\|}{\min_{1 \leqslant i \leqslant m} -f_i(\hat{x}(\Pi_S(x)))} \leqslant \hat{r} < \infty.$$

From (3.37), we have

$$\begin{aligned}
dist(x, S) &= \|x - \Pi_S(x)\| \\
&\leqslant \frac{\|x - \hat{x}(\Pi_S(x))\| - \|x - \Pi_S(x)\|}{\min_{1 \leqslant i \leqslant m} -f_i(\hat{x}(\Pi_S(x)))} \|F(x)_+\|_\infty \\
&\leqslant \frac{\|x - \Pi_S(x)\| + \|\Pi_S(x) - \hat{x}(\Pi_S(x))\| - \|x - \Pi_S(x)\|}{\min_{1 \leqslant i \leqslant m} -f_i(\hat{x}(\Pi_S(x)))} \|F(x)_+\|_\infty \\
&= \frac{\|\Pi_S(x) - \hat{x}(\Pi_S(x))\|}{\min_{1 \leqslant i \leqslant m} -f_i(\hat{x}(\Pi_S(x)))} \|F(x)_+\|_\infty \\
&\leqslant \hat{r} \|F(x)_+\|_\infty.
\end{aligned}$$

□

On the other hand, we can prove Theorem 3.4.1 by using Corollary 3.3.1.

Alternative proof of Theorem 3.4.1

Let $f(x) := \max_{1 \leqslant i \leqslant m} f_i(x) \quad$ for $x \in \mathbb{R}^n$. Then f is a real-valued convex function on $\mathbb{R}^n$ and $S = f^{-1}(-\infty, 0]$. We claim that the Strong Slater CQ implies that $0 \notin \overline{\partial f(f^{-1}(0))}$. Hence, by Corollary 3.3.1, there exists an $\delta > 0$ such that

$$dist(x, S) \leqslant \delta f(x)_+ = \delta \|F(x)_+\|_\infty \quad \text{for all } x \in \mathbb{R}^n.$$

Suppose on the contrary that $0 \in \overline{\partial f(f^{-1}(0))}$. Then there exist sequences $\{ x^k \} \subseteq f^{-1}(0)$ and $\{ a^k \}$ with each $a^k \in \partial f(x^k)$ such that

$$\lim_{k \to \infty} a^k = 0.$$

Note that $f^{-1}(0) \subseteq \partial S$, it follows from Lemma 3.4.1 that there exists a sequence $\{ z^k \} \subseteq S^0$ satisfying

$$\|x^k - z^k\| \leqslant \hat{r} \min_{1 \leqslant i \leqslant m} -f_i(z^k)$$

or equivalently

$$-\hat{r}^{-1}\|x^k - z^k\| \geqslant f(z^k).$$

On the other hand, since $a^k \in \partial f(x^k)$, we have

$$\begin{aligned} -\hat{r}^{-1}\|x^k - z^k\| &\geqslant f(z^k) \\ &= f(z^k) - f(x^k) \\ &\geqslant \langle a^k, z^k - x^k \rangle \\ &\geqslant -\|a^k\|\|z^k - x^k\|. \end{aligned}$$

Hence, $\|a^k\| \geqslant \hat{r}^{-1}$ for all $k \in \mathbb{N}$. Let k go to infinity, we get a contradiction that $0 > \hat{r}^{-1}$.

□

Note that we do not know the value of δ explicitly.

3.4.2 Constrained Case

In this subsection, we consider X to be a real reflexive Banach space and C to be a non-empty closed convex subset of X.

Assumption 1

There exists an $\bar{u} \in C^\infty$ and a constant $r > 0$ such that

$$f_{i\infty}(\bar{u}) \leqslant -r^{-1} \quad \text{for all } i = 1, \ldots, m.$$

Proposition 3.4.1 *If Assumption 1 holds, then*

(a) S_C is unbounded,

(b) there exists an $y \in C$ such that $f_i(y) < 0$ for all $i \in \{1, \ldots, m\}$.

<u>**Proof**</u>

<u>(a)</u>

Suppose S_C is bounded and let $x_0 \in S_C \subseteq C$. Note that $\bar{u} \neq 0$ and $x_0 + \lambda\bar{u} \in C$ for all $\lambda > 0$ as $\bar{u} \in C^\infty$. Therefore $x_0 + \bar{\lambda}\bar{u} \in C \backslash S_C$ for some $\bar{\lambda} > 0$ large enough. Hence, by the definition of S_C, $f_i(x_0 + \bar{\lambda}\bar{u}) > 0$ for some $i \in \{1, \ldots, m\}$ and so

$$0 > f_{i\infty}(\bar{u}) = \sup_{\lambda>0} \frac{f_i(x_0 + \lambda\bar{u}) - f_i(x_0)}{\lambda} \geqslant \frac{f_i(x_0 + \bar{\lambda}\bar{u}) - f_i(x_0)}{\bar{\lambda}} > 0.$$

We get a contradiction.

<u>(b)</u>

By Assumption 1, $f_i(x + \lambda\bar{u}) \leqslant f_i(x) - \lambda r^{-1}$ for all $i \in \{1, \ldots m\}$, $x \in X$ and $\lambda > 0$. Therefore, for any $x \in C$, we can take $y = x + \lambda\bar{u} \in C$ for some $\lambda > 0$ large enough such that (b) holds.

□

Theorem 3.4.4 (Deng [7]) *If Assumption 1 holds with $\bar{u}$ being a unit vector, then for any p with $1 \leqslant p \leqslant \infty$, we have*

$$dist(x, S_C) \leqslant r\|F(x)_+\|_p \quad \textit{for all } x \in C,$$

where r is defined as in Assumption 1.

<u>**Proof**</u>

Let $f : X \to \mathbb{R}$ be a continuous convex function on X defined by

$$x \mapsto \max_{1 \leqslant i \leqslant m} f_i(x).$$

Then $S_C = \{x \in C \mid f(x) \leqslant 0\}$. Fix an $x \in C \backslash S_C$ and let $\bar{x} = \Pi_{S_C}(x)$, we have $0 \in \partial\|\bar{x} - x\| + N(\bar{x}; S_C)$ as $\bar{x}$ is a global minimizer of the map $y \mapsto \|y - z\| + I_{S_C}(y)$. Therefore, we can write $0 = v_1 + v_2$ for some $v_1 \in \partial\|\bar{x} - x\|$ and $v_2 \in N(\bar{x}; S_C)$. Also, by Theorem 2.2.11, we know that $\|\bar{x} - x\| = \langle -v_1, x - \bar{x}\rangle$ and $\|v_1\| = 1$.

Note that $f(\bar{x}) = 0$, so $\bar{x}$ is not a global minimizer of f by Proposition 3.4.1 and hence by Lemma 3.3.2, we have $N(\bar{x}; S_C) = cone(\partial f(\bar{x})) + N(\bar{x}; C)$. Therefore, we can write $-v_1 = v_2 = \lambda u_1 + u_2$ where $\lambda \geqslant 0$, $u_1 \in \partial f(\bar{x})$ and $u_2 \in N(\bar{x}; C)$. By Theorem 2.2.16, we can write $u_1 = \sum_{i \in T(\bar{x})} \alpha_i w_i$, where $\alpha_i \geqslant 0$ with $\sum_{i \in T(\bar{x})} \alpha_i = 1$ and $w_i \in \partial f_i(\bar{x})$. So

$$\begin{aligned}
1 = \| -v_1 \| &\geqslant \langle -v_1, -\bar{u} \rangle = \lambda \langle u_1, -\bar{u} \rangle + \langle u_2, -\bar{u} \rangle \\
&= \lambda \langle u_1, -\bar{u} \rangle - \langle u_2, \bar{x} + \bar{u} - \bar{x} \rangle \geqslant \lambda \langle u_1, -\bar{u} \rangle \\
&= \lambda \left(\sum_{i \in T(\bar{x})} \alpha_i \langle w_i, -\bar{u} \rangle \right). \qquad (3.38)
\end{aligned}$$

On the other hand, by Assumption 1, for any $i \in T(\bar{x})$,

$$\begin{aligned}
-r^{-1} &\geqslant f_{i\infty}(\bar{u}) = \sup_{x \in X} \{ f_i(x + \bar{u}) - f_i(x) \} \\
&\geqslant \sup_{x \in X} \sup_{v \in \partial f_i(x)} \langle v, \bar{u} \rangle \geqslant \sup_{v \in \partial f_i(\bar{x})} \langle v, \bar{u} \rangle \\
&\geqslant \langle w_i, \bar{u} \rangle. \qquad (3.39)
\end{aligned}$$

By (3.38) and (3.39), we have

$$\lambda r^{-1} = \sum_{i \in T(\bar{x})} \lambda \alpha_i r^{-1} \leqslant \lambda \left(\sum_{i \in T(\bar{x})} \alpha_i \langle w_i, -\bar{u} \rangle \right) \leqslant 1.$$

Hence, $\lambda \leqslant r$ and

$$\begin{aligned}
dist(x, S_C) = \|x - \bar{x}\| &= \langle -v_1, x - \bar{x} \rangle = \lambda \langle u_1, x - \bar{x} \rangle + \langle u_2, x - \bar{x} \rangle \\
&\leqslant \lambda \langle u_1, x - \bar{x} \rangle = \lambda \sum_{i \in T(\bar{x})} \alpha_i \langle w_i, x - \bar{x} \rangle \leqslant \lambda \sum_{i \in T(\bar{x})} \alpha_i (f_i(x) - f_i(\bar{x})) \\
&\leqslant \lambda \sum_{i \in T(\bar{x})} \alpha_i \max\{ f_i(x), 0 \} \leqslant r \sum_{i \in T(\bar{x})} \alpha_i \max\{ f_i(x), 0 \} \\
&\leqslant r \left(\sum_{i \in T(\bar{x})} \alpha_i^q \right)^{\frac{1}{q}} \left(\sum_{i \in T(\bar{x})} (\max\{ f_i(x), 0 \})^p \right)^{\frac{1}{p}} \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1 \\
&\leqslant r \| F(x)_+ \|_p.
\end{aligned}$$

□

Chapter 4

Some Recent Results on Certain Proper Efficient Points

4.1 Scalarization of Henig Proper Efficient Points

Throughout this section, we use E to denote a real normed vector space, $C \subseteq E$ is a closed convex cone with a base Θ. That is, Θ is a convex subset of C with $0 \notin \overline{\Theta}$ and $C = cone(\Theta)$, where $\overline{\Theta}$ is the closure of Θ. Note that, C must be pointed whenever it has a base, for otherwise one may take a non-zero element x in $C \cap -C$ and so there exists $\lambda_i > 0$, $b_i \in \Theta$ for $i = 1, 2$, such that $x = \lambda_1 b_1$ and $-x = \lambda_2 b_2$. But then

$$0 = \frac{x + (-x)}{\lambda_1 + \lambda_2} = \frac{\lambda_1 b_1 + \lambda_2 b_2}{\lambda_1 + \lambda_2} \in \Theta,$$

which contradicting the definition of Θ. Let $\leqslant_C$ be the order induced by C: for $x, y \in E$, $x \leqslant_C y$ if and only if $y - x \in C$.

Definition 4.1.1 *Let A be a non-empty subset of E, we say that*

(a) $x \in A$ to be an efficient point of A with respect to C, written as $x \in E(A, C)$ if

$$(A - x) \cap -C = \{0\};$$

that is to say that there does not exist $a \in A\backslash\{x\}$ such that $a \leqslant x$.

(b) $x \in A$ to be a Henig proper point of A with respect to Θ, written as $x \in HE(A, \Theta)$, if there exists an $\epsilon > 0$ such that

$$\overline{cone(A-x)} \cap -\overline{cone(\Theta+\epsilon\mathbb{B})} = \{0\},$$

where $\mathbb{B}$ is the closed unit ball in E;

(c) $x \in A$ to be a super efficient point of A with respect to C, written as $x \in SE(A, C)$, if there exists an $M > 0$ such that

$$\overline{cone(A-x)} \cap (\mathbb{B}-C) \subseteq M\mathbb{B}.$$

Remark: $HE(A, \Theta)$ may depend on the choice of Θ. However, when Θ is bounded, we will see in the following that $HE(A, \Theta) = SE(A, C)$ and hence $HE(A, \Theta)$ is independent of the choice of Θ.

In the following of this section, we shall first study relationship between the various concepts of efficient points introduced above and then characterize the Heing proper efficient point by a family of continuous monotone Minkowski functionals and a family of continuous norms. These results are taken from [5] and [21].

4.1.1 Preliminaries

Let $\delta := \inf\{\,\|\theta\| \mid \theta \in \overline{\Theta}\,\}$. Then $\delta > 0$ because $0 \notin \overline{\Theta}$. For each $n \in \mathbb{N}$, define $S_n := \overline{cone(\Theta + \frac{\delta}{2n}\mathbb{B})}$ and $S_n^0 := cone(\Theta + \frac{\delta}{2n}\mathbb{B}^0)$, where $\mathbb{B}^0$ is the open unit ball in E. Note that

(a) $0 \notin \overline{\Theta + \frac{\delta}{2n}\mathbb{B}}$;

(b) $x \in HE(A,\Theta)$ if and only if $\overline{cone(A-x)} \cap -S_n = \{0\}$ for some $n \in \mathbb{N}$.

Proposition 4.1.1 *For any $n \in \mathbb{N}$, S_n is a closed, convex and pointed cone with a base $(\overline{\Theta + \frac{\delta}{2n}\mathbb{B}})$. That is, $S_n = cone(\overline{\Theta + \frac{\delta}{2n}\mathbb{B}})$.*

Proof

Clearly, S_n is closed, and convex by the convexity of $\Theta + \frac{\delta}{2n}\mathbb{B}$. Now, it remains to prove the equality. Let $x \in cone(\overline{\Theta + \frac{\delta}{2n}\mathbb{B}})$. Then $x = \lambda \lim_{k\to\infty} x_k$, where $\lambda \geqslant 0$ and $x_k \in \Theta + \frac{\delta}{2n}\mathbb{B}$ for all $k \in \mathbb{N}$. Hence, $x = \lim_{k\to\infty} \lambda x_k \in \overline{cone(\Theta + \frac{\delta}{2n}\mathbb{B})}$ showing that $cone(\overline{\Theta + \frac{\delta}{2n}\mathbb{B}}) \subseteq S_n$. For the reverse inclusion, let $y \in S_n = \overline{cone(\Theta + \frac{\delta}{2n}\mathbb{B})}$. Then $y = \lim_{k\to\infty} \alpha_k y_k$, where $\alpha_k \geqslant 0$ and $y_k \in \Theta + \frac{\delta}{2n}\mathbb{B}$ for all $n \in \mathbb{N}$. We claim that, taking a subsequence if necessary, α_k converges in $\mathbb{R}$. Suppose not, α_k will tend to positive infinity. Then $y_k \to 0$, contradicting the fact that $0 \notin \overline{\Theta + \frac{\delta}{2n}\mathbb{B}}$.

Case1 $\alpha_k \to 0$.

Write $\alpha_k y_k = \alpha_k \theta_k + \alpha_k \frac{\delta}{2n} b_k$, where $\theta_k \in \Theta$ and $b_k \in \mathbb{B}$ for all $k \in \mathbb{N}$. As $\alpha_k \theta_k \in C$ for all k and $\lim_{k\to\infty} \alpha_k \frac{\delta}{2n} = 0$. Therefore,

$$y \in \cap_{\epsilon>0}(C + \epsilon\mathbb{B}) = \overline{C} = C \subseteq cone\left(\overline{\Theta + \frac{\delta}{2n}\mathbb{B}}\right).$$

Case2 $\alpha_k \to \alpha > 0$.

Then $\lim_{k\to\infty} y_k = \frac{y}{\alpha} \in \overline{\Theta + \frac{\delta}{2n}\mathbb{B}}$. Therefore $y \in cone(\overline{\Theta + \frac{\delta}{2n}\mathbb{B}})$.

□

Lemma 4.1.1 *Let A be a subset of E, $x \in A$ and $n \in \mathbb{N}$. Suppose that $(A-x) \cap -S_n^0 = \{0\}$. Then $\overline{cone(A-x)} \cap -S_{n+1} = \{0\}$ and so $x \in HE(A,\Theta)$.*

Proof

By supposition, $cone(A-x) \cap -S_n^0 = \{0\}$, that is

$$cone(A-x) \cap -(\Theta + \frac{\delta}{2n}\mathbb{B}^0) = \emptyset. \tag{4.1}$$

Let $r := \frac{1}{2}\left(\frac{\delta}{2n} - \frac{\delta}{2(n+1)}\right)$, we have $\frac{\delta}{2(n+1)}\mathbb{B}^0 + r\mathbb{B} + r\mathbb{B} \subseteq \frac{\delta}{2n}\mathbb{B}^0$. This and (4.1) imply that

$$(cone(A-x) + r\mathbb{B}) \cap -\left(\Theta + \frac{\delta}{2(n+1)}\mathbb{B}^0 + r\mathbb{B}\right) = \emptyset.$$

Since

$$\overline{cone(A-x)} \subseteq cone(A-x) + r\mathbb{B}$$

and $\overline{\Theta + \frac{\delta}{2(n+1)}\mathbb{B}} \subseteq \Theta + \frac{\delta}{2(n+1)}\mathbb{B} + \frac{r}{2}\mathbb{B} \subseteq \Theta + \frac{\delta}{2(n+1)}\mathbb{B}^0 + r\mathbb{B}$, we have

$$\overline{cone(A-x)} \cap -\overline{\left(\Theta + \frac{\delta}{2(n+1)}\mathbb{B}\right)} = \emptyset$$

and so

$$\overline{cone(A-x)} \cap -S_{n+1} = \{\, 0 \,\}.$$

□

Lemma 4.1.2 *If $x \in SE(A,C)$, then there exists $M > 0$ such that for each $a \in A$ and $y \in E$, the following implication holds*

$$a - x \leqslant_C y \Rightarrow \|a-x\| \leqslant M\|y\|. \tag{4.2}$$

Proof

Let $x \in SE(A,C)$, so there exists $M > 0$ such that $\overline{cone(A-x)} \cap (B-C) \subseteq M\mathbb{B}$. Suppose $a - x \leqslant_C y$. If $y = 0$, then $a - x \in (A-x) \cap -C$. Hence,

$$\lambda(a-x) \in cone(A-x) \cap -C \subseteq M\mathbb{B}$$

for all $\lambda \geqslant 0$. Therefore, we must have $a - x = 0$ and (4.2) holds in the case when $y = 0$.

If $y \neq 0$, then $(a-x)/\|y\| \leqslant_C y/\|y\|$. Hence, there exists an $c_0 \in C$ such that $(a-x)/\|y\| = y/\|y\| - c_0 \in \mathbb{B} - C$. Therefore, $(a-x)/\|y\| \in cone(A-x) \cap (\mathbb{B} - C) \subseteq M\mathbb{B}$ an so $\|a-x\| \leqslant M\|y\|$.

□

The converse of Lemma 4.1.2 is also true, as the following lemma shows.

Lemma 4.1.3 *Let $x \in A$. If there exists $M > 0$ such that for each $a \in A$ and $y \in E$, $a - x \leqslant_C y \Rightarrow \|a - x\| \leqslant M\|y\|$, then $x \in SE(A, C)$.*

Proof

Let x be given as above. Then we have $cone(A-x) \cap (\mathbb{B}-C) \subseteq M\mathbb{B}$. Fix $\epsilon > 0$ and let $z \in \overline{cone(A - x)} \cap (\mathbb{B}-C)$. Write $z = u + \epsilon b_0 = b - c_0$ for some $u \in cone(A-x)$, b and $b_0 \in \mathbb{B}$, $c_0 \in C$. $u = b - \epsilon b_0 - c_0 \in (1+\epsilon)\mathbb{B} - C \subseteq (1+\epsilon)(\mathbb{B} - C)$. Therefore, $u \in (1 + \epsilon)(cone(A - x) \cap (\mathbb{B} - C)) \subseteq (1 + \epsilon)M\mathbb{B}$. Hence, $\|z\| \leqslant \|u\| + \|\epsilon b_0\| \leqslant (1 + \epsilon)M + \epsilon$. It follows that $\overline{cone(A - x)} \cap (\mathbb{B} - C) \subseteq ((1 + \epsilon)M + \epsilon)\mathbb{B}$ and so $x \in SE(A, C)$.

□

Proposition 4.1.2 *If C has a base Θ, then $SE(A, C) \subseteq HE(A, \Theta)$.*

Proof

Let $x \in SE(A, C)$ and $\epsilon > 0$ with $\epsilon < \frac{\delta}{1+M}$. We claim that $cone(A-x) \cap (\epsilon\mathbb{B}-\Theta) = \emptyset$. Suppose the contrary that there exists an $z = \lambda(a - x) = \epsilon b - \theta \leqslant_C \epsilon b$, for some $b \in \mathbb{B}$, $\theta \in \Theta$, $a \in A$ and $\lambda > 0$. Then $\lambda^{-1}\|z\| \leqslant \lambda^{-1}M\epsilon$ by Lemma 4.1.2. Hence,

$$\lambda^{-1}(\delta - \epsilon) \leqslant \lambda^{-1}(\|\theta\| - \|\epsilon b\|) \leqslant \lambda^{-1}(\|\theta - \epsilon b\|) = \lambda^{-1}\|z\| \leqslant \lambda^{-1}M\epsilon.$$

We get $\frac{\delta}{1+M} \leqslant \epsilon$ contradicting the choice of ϵ. Therefore,

$$cone(A - x) \cap (\epsilon\mathbb{B} - \Theta) = \emptyset$$

and this implies that

$$cone(A - x) \cap -cone(\epsilon\mathbb{B} + \Theta) = \{0\} \quad \text{and} \quad (A - x) \cap -S_n^0 = \{0\}$$

for some $n \in \mathbb{N}$. By Lemma 4.1.1, $x \in HE(A, \Theta)$.

□

Proposition 4.1.3 *Suppose that C has a bounded base Θ. Then $HE(A, \Theta) = SE(A, C)$. Consequently, $HE(A, \Theta)$ does not depend on the choice of Θ whenever C has a bounded base.*

Proof

By Proposition 4.1.2, it suffices to prove that $HE(A, \Theta) \subseteq SE(A, C)$. Let $x \in HE(A, \Theta)$. By definition, $\overline{cone(A - x)} \cap -\overline{cone(\Theta + \frac{\delta}{2n}\mathbb{B})} = \{0\}$ for some $n \in \mathbb{N}$ where $\delta := \inf\{\, \|\theta\| \mid \theta \in \overline{\Theta}\,\}$. Therefore,

$$cone(A - x) \cap \left(\frac{\delta}{2n}\mathbb{B} - \Theta\right) = \emptyset. \tag{4.3}$$

Let $y \in E$ be such that $a - x \leqslant_C y$ for some $a \in A$ and write $a - x = y - \lambda\theta$ for some $\lambda \geqslant 0$ and $\theta \in \Theta$. If $\lambda > 0$, then $\|\lambda^{-1}y\| > \frac{\delta}{2n}$ by (4.3), so $y \neq 0$ with $\|y\|^{-1} \leqslant 2n/\lambda\delta$. Set $m := \sup\{\, \|\theta\| \mid \theta \in \Theta\,\} < \infty$, and $M = 1 + 2nm/\delta$. Then

$$\begin{aligned} \|a - x\| \leqslant \|y\| + \lambda\|\theta\| &\leqslant \|y\| + \lambda m \\ &= \|y\|(1 + \lambda m\|y\|^{-1}) \\ &\leqslant M\|y\|. \end{aligned}$$

Obviously, the above inequality also holds for $\lambda = 0$. By Lemma 4.1.3, $x \in SE(A, C)$.

□

4.1.2 Scalarization by Monotone Minkowski Functionals

In the following, we fix an $\theta_0 \in \Theta$. Let $b \in \mathbb{B}^0$ and $n \in \mathbb{N}$. Then since S_n contains the open set $\theta_0 + \frac{\delta}{2n}\mathbb{B}_n^0$ containing $\theta_0 + \frac{\delta}{2n}b$. Hence $\theta_0 + \frac{\delta}{2n}b - S_n$ is a neighborhood of 0. Let $P_b^{(n)}$ be the Minkowski functional of the closed convex neighborhood $\theta_0 + \frac{\delta}{2n}b - S_n$ of the origin. That is,

$$P_b^{(n)}(x) = \inf\{\, t > 0 \mid x \in t(\theta_0 + \frac{\delta}{2n}b - S_n)\,\} \quad \text{for } x \in E.$$

It is known that $P_b^{(n)}$ is positively homogeneous, subadditive and continuous. We have

$$x \in (P_b^{(n)}(x) + \epsilon)(\theta_0 + \frac{\delta}{2n} b - S_n) \quad \text{for any } \epsilon > 0. \tag{4.4}$$

Proposition 4.1.4 *(a) For each $n \in \mathbb{N}$ and $b \in \mathbb{B}^0$, $\{ x \in E \mid P_b^{(n)}(x) = 0 \} = -S_n$.*

(b) For each $n \in \mathbb{N}$ and $b \in \mathbb{B}^0$, $P_b^{(n)}(x) \leqslant P_b^{(n)}(y)$ whenever $x \leqslant_C y$.

(c) For each $(b, x) \in \mathbb{B}^0 \times E$, let $\Phi_n(b, x) := P_b^{(n)}(x)$. Then for any $0 < \epsilon < 1$ and $r > 0$, $\Phi_n(b, x)$ is a Lipschitz function on $(1-\epsilon)\mathbb{B} \times r\mathbb{B}$. Hence, Φ_n is continuous on $\mathbb{B}^0 \times E$.

Proof

(a)

Let $x_0 \in \{ x \in E \mid P_b^{(n)}(x) = 0 \}$ and write ω for $\theta_0 + \frac{\delta}{2n} b (\in S_n)$. From (4.4), $x_0 \in \frac{1}{k}(\omega - S_n) = \frac{1}{k}\omega - S_n$ for all $k \in \mathbb{N}$. Then $x_0 \in \overline{-S_n} = -S_n$ and so $\{ x \in E \mid P_b^{(n)}(x) = 0 \} \subseteq -S_n$. Conversely, note that $-S_n \subseteq \omega - S_n$ (because any $-s \in -S_n$ can be expressed as $\omega - (\omega + s)$ with $\omega + s \in S_n$.) Consequently, if $z \in S_n$, then $-kz \in -S_n \subseteq \omega - S_n$ for all $k \in \mathbb{N}$ and so $P_b^{(n)}(-z) = 0$. That is, $-S_n \subseteq \{ x \in E \mid P_b^{(n)}(x) = 0 \}$.

(b)

Suppose $x \leqslant_C y$. Then there exist $t \geqslant 0$ and $\theta \in \Theta$ such that $x = y - t\theta$. From (4.4), for any $\epsilon > 0$,

$$x \in (P_b^{(n)}(y) + \epsilon)(\theta_0 + \frac{\delta}{2n} b - S_n) - t\theta \subseteq (P_b^{(n)}(y) + \epsilon)(\theta_0 + \frac{\delta}{2n} b - S_n).$$

This implies that $P_b^{(n)}(x) \leqslant P_b^{(n)}(y)$ since $\epsilon > 0$ is arbitrary.

(c)

For any $0 < \epsilon < 1$, $r > 0$, $\alpha > 0$, (a, x) and $(b, y) \in (1-\epsilon)\mathbb{B} \times r\mathbb{B}$,

$$x \in (\Phi_n(a, x) + \alpha)(\theta_0 + \frac{\delta}{2n} a - S_n). \tag{4.5}$$

Since $\|b\| \leqslant 1-\epsilon$ and S_n contains $\theta_0 + \frac{\delta}{2n}\mathbb{B}$, we have

$$\frac{\delta\epsilon}{2n}\mathbb{B} \subseteq \frac{\delta}{2n}(b-\mathbb{B}) \subseteq \theta_0 + \frac{\delta}{2n}b - S_n$$

and so

$$\begin{aligned} &\frac{\delta\epsilon(y-x)}{2n\|y-x\|} \in \theta_0 + \frac{\delta}{2n}b - S_n \\ \text{and} \quad &\frac{\delta\epsilon(a-b)}{2n\|a-b\|} \in \theta_0 + \frac{\delta}{2n}b - S_n. \end{aligned}$$

This implies that

$$\begin{aligned} &y \in x + \frac{2n\|y-x\|}{\delta\epsilon}(\theta_0 + \frac{\delta}{2n}b - S_n) \\ \text{and} \quad &a \in b + \frac{2n\|a-b\|}{\delta\epsilon}(\theta_0 + \frac{\delta}{2n}b - S_n) \end{aligned} \tag{4.6}$$

(the relations are respectively trivial if $y = x$ or $a = b$). From the second relation of (4.6), one has

$$\theta_0 + \frac{\delta}{2n}a \in \theta_0 + \left(\frac{\delta}{2n}b + \frac{\|a-b\|}{\epsilon}(\theta_0 + \frac{\delta}{2n}b - S_n)\right) \subseteq (1+\frac{\|a-b\|}{\epsilon})(\theta_0 + \frac{\delta}{2n}b - S_n).$$

It follows from (4.5),

$$\begin{aligned} x &\in (\Phi_n(a,x)+\alpha)\left((1+\frac{\|a-b\|}{\epsilon})(\theta_0 + \frac{\delta}{2n}b - S_n) - S_n\right) \\ &\subseteq (\Phi_n(a,x)+\alpha)(1+\frac{\|a-b\|}{\epsilon})(\theta_0 + \frac{\delta}{2n}b - S_n). \end{aligned}$$

This and (4.6) imply that

$$y \in \left((\Phi_n(a,x)+\alpha)(1+\frac{\|a-b\|}{\epsilon}) + \frac{2n\|y-x\|}{\delta\epsilon}\right)(\theta_0 + \frac{\delta}{2n}b - S_n).$$

Since α is arbitrary,

$$\Phi_n(b,y) \leqslant \Phi_n(a,x)\left(1+\frac{\|a-b\|}{\epsilon}\right) + \frac{2n\|y-x\|}{\delta\epsilon}. \tag{4.7}$$

Setting $(a,x) = (0,0)$, we have

$$\Phi_n(b,y) \leqslant \frac{2nr}{\delta\epsilon}, \tag{4.8}$$

for any $(b, y) \in (1-\epsilon)\mathbb{B} \times r\mathbb{B}$.

Let $M_n := \max\{\frac{2nr}{\delta\epsilon^2}, \frac{2n}{\delta\epsilon}\}$. By (4.7) and (4.8), for any (a, x), $(b, y) \in (1-\epsilon)\mathbb{B} \times r\mathbb{B}$,

$$\begin{aligned}\Phi_n(b, y) - \Phi_n(a, x) &\leqslant \Phi_n(a, x)\frac{\|a-b\|}{\epsilon} + \frac{2n}{\delta\epsilon}\|y-x\| \\ &\leqslant (\frac{2nr}{\delta\epsilon^2})\|a-b\| + \frac{2n}{\delta\epsilon}\|y-x\| \\ &\leqslant M_n(\|a-b\| + \|y-x\|).\end{aligned}$$

By symmetry, we also have

$$\Phi_n(a, x) - \Phi_n(b, y) \leqslant M_n(\|a-b\| + \|y-x\|).$$

Hence, $|\Phi_n(a, x) - \Phi_n(b, y)| \leqslant M_n(\|a-b\| + \|y-x\|)$. This completes the proof.

□

Theorem 4.1.1 (Zheng [21]) *Let $A \subseteq E$. Then $x \in HE(A, \Theta)$ if and only if $x \in A$ and there exist $b \in \mathbb{B}^0$, $n > \|x\|$ such that*

$$P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0) = \inf\{\, P_b^{(n)}(y + \frac{2n^2}{\delta}\theta_0) \mid y \in A\,\}.$$

(Note that n can be arbitrarily large for the 'if' part, see the proof below.)

<u>**Proof**</u>

Suppose $x \in HE(A, \Theta)$. Then there exists $n > \|x\|$ sufficiently large such that $\overline{cone(A-x)} \cap -S_n = \{\,0\,\}$ and this implies that for any $y \in A\backslash\{\,x\,\}$,

$$y + \frac{2n^2}{\delta}\theta_0 \notin \frac{2n^2}{\delta}\left(\theta_0 + \frac{\delta}{2n}(\frac{x}{n}) - S_n\right).$$

Therefore, for any $y \in A\backslash\{\,x\,\}$,

$$P_{\frac{x}{n}}^{(n)}(y + \frac{2n^2}{\delta}\theta_0) \geqslant \frac{2n^2}{\delta}.$$

As $x + \frac{2n^2}{\delta}\theta_0 \in \frac{2n^2}{\delta}\left(\theta_0 + \frac{\delta}{2n}(\frac{x}{n}) - S_n\right)$,

$$P_{\frac{x}{n}}^{(n)}(x + \frac{2n^2}{\delta}\theta_0) \leqslant \frac{2n^2}{\delta}.$$

Hence,

$$P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0) = \inf\{ P_b^{(n)}(y + \frac{2n^2}{\delta}\theta_0) \mid y \in A \},$$

where $b := x/n \in \mathbb{B}^0$.

Conversely, suppose that there are $b \in \mathbb{B}^0$ and $n \in \mathbb{N}$ with $n > \|x\|$ such that

$$P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0) = \inf\{ P_b^{(n)}(z + \frac{2n^2}{\delta}\theta_0) \mid z \in A \}. \tag{4.9}$$

For any $y \in (x - S_n^0)\backslash\{ x \}$, there exist $t > 0$, $\theta \in \Theta$ and $a \in \mathbb{B}^0$ such that $y = x - t(\theta + \frac{\delta}{2n}a)$. Set $r = (1 - \|a\|)/\|\theta_0 + \frac{\delta}{2n}b\| > 0$ and $b_1 = a - r(\theta_0 + \frac{\delta}{2n}b)$. Then,

$$b_1 \in \mathbb{B} \quad \text{and} \quad \frac{\delta}{2n}a = \frac{\delta}{2n}b_1 + \frac{\delta r}{2n}(\theta_0 + \frac{\delta}{2n}b). \tag{4.10}$$

For any $\epsilon > 0$, we have by (4.4) that

$$y + \frac{2n^2}{\delta}\theta_0 \in (P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0) + \epsilon)(\theta_0 + \frac{\delta}{2n}b - S_n) - t(\theta + \frac{\delta}{2n}a). \tag{4.11}$$

Since $n \in \mathbb{N}$ and $n > \|x\|$, $x + \frac{2n^2}{\delta}\theta_0 \neq 0$ and $x + \frac{2n^2}{\delta}\theta_0 = \frac{2n^2}{\delta}\left(\theta_0 + \frac{\delta}{2n}(\frac{x}{n})\right) \in S_n$. As S_n is pointed, $x + \frac{2n^2}{\delta}\theta_0 \notin -S_n$ and so $P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0) > 0$ by (a) of Proposition 4.1.4. Hence, there are $0 < r_1$ and $0 < r_2 < P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0)$ such that $r_1 + r_2 = \frac{t\delta r}{2n}$. From (4.10), we have

$$\begin{aligned} t(\theta + \frac{\delta}{2n}a) &= t(\theta + \frac{\delta}{2n}b_1) + r_1(\theta_0 + \frac{\delta}{2n}b) + r_2(\theta_0 + \frac{\delta}{2n}b) \\ &\in S_n + S_n + r_2(\theta_0 + \frac{\delta}{2n}b) = r_2(\theta_0 + \frac{\delta}{2n}b + S_n). \end{aligned}$$

This and (4.11) imply that

$$y + \frac{2n^2}{\delta}\theta_0 \in \left(P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0) + \epsilon - r_2\right)(\theta_0 + \frac{\delta}{2n}b + S_n).$$

Since ϵ is arbitrary,

$$\begin{aligned} P_b^{(n)}(y + \frac{2n^2}{\delta}\theta_0) &\leqslant P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0) - r_2 \\ &< P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0). \end{aligned}$$

From (4.9), we have $y \notin A$. Therefore, $y \notin A$ whenever $y \in (x - S_n^0)\backslash\{x\}$, which implies that $(A - x) \cap -S_n^0 = \{0\}$. By lemma 4.1.1, $x \in HE(A, \Theta)$.

□

By Proposition 4.1.3, we have the following theorem immediately.

Theorem 4.1.2 (Zheng [21]) *Let $A \subseteq E$ and Θ be bounded. Then $x \in SE(A, C)$ if and only if $x \in A$ and there exist $b \in \mathbb{B}^0$, $n > \|x\|$ such that*

$$P_b^{(n)}(x + \frac{2n^2}{\delta}\theta_0) = \inf\{ P_b^{(n)}(y + \frac{2n^2}{\delta}\theta_0) \mid y \in A \}.$$

4.1.3 Scalarization by Continuous Norms

For each $b \in \mathbb{B}^0$ and $n \in \mathbb{N}$, let

$$U(n, b) := (-\theta_0 - \frac{\delta}{2n}b + S_n) \cap (\theta_0 + \frac{\delta}{2n}b - S_n)$$

and define $Q_b^{(n)} : E \to \mathbb{R}^+$ such that

$$Q_b^{(n)}(x) = \inf\{ t > 0 \mid x \in tU(n, b) \} \quad \text{for } x \in E.$$

Proposition 4.1.5 *For any $b \in \mathbb{B}^0$ and $n \in \mathbb{N}$, $Q_b^{(n)}$ is a continuous norm on E and if Θ is bounded, then $Q_b^{(n)}$ is an equivalent norm on E.*

Proof

Since $b \in \mathbb{B}^0$, $\theta_0 + \frac{\delta}{2n}b - S_n$ and $-\theta_0 - \frac{\delta}{2n}b + S_n$ are neighborhoods of 0 in E as in the previous section. Moreover, $U(n, b)$ is balanced, closed and convex. Therefore, $Q_b^{(n)}$ is a continuous seminorm on E. To prove that $Q_b^{(n)}$ is a norm, it suffices to show that for any $x \in E\backslash\{0\}$, $Q_b^{(n)}(x) > 0$. Suppose that $x_0 \in E$ such that $Q_b^{(n)}(x_0) = 0$. Since $Q_b^{(n)}$ is a seminorm on E, $Q_b^{(n)}(tx_0) = 0$ for all $t \in \mathbb{R}$. Hence, for any $r \in \mathbb{R}$,

$$rx_0 \in \{ x \in E \mid Q_b^{(n)}(x) \leqslant 1 \} = U(n, b).$$

It follows that for any $r > 0$,

$$x_0 \in \left(\frac{1}{r}(\theta_0 + \frac{\delta}{2n}b) - S_n\right) \cap \left(\frac{1}{r}(-\theta_0 - \frac{\delta}{2n}b) + S_n\right).$$

Let $r \to \infty$, we have $x_0 \in S_n \cap -S_n = \{0\}$ as S_n is closed and pointed. Therefore, $Q_b^{(n)}$ is a norm. To prove that $Q_b^{(n)}$ is an equivalent norm if Θ is bounded, it suffices to show that $U(n,b)$ is bounded. For any $x \in U(n,b)$, there exist $\theta_i \in \Theta$, $b_i \in \mathbb{B}$ and $t_i \geqslant 0$ for $i = 1, 2$ such that

$$x = \theta_0 + \frac{\delta}{2n}b - t_1(\theta_1 + \frac{\delta}{1.5}b_1) = -\theta_0 - \frac{\delta}{2n}b + t_2(\theta_2 + \frac{\delta}{1.5}b_2) \tag{4.12}$$

as $S_n = cone\overline{(\Theta + \frac{\delta}{2n}\mathbb{B})} \subseteq cone(\Theta + \frac{\delta}{1.5}\mathbb{B})$.

Hence,

$$\begin{aligned}
2(\|\theta_0\| + \frac{\delta}{2n}) &\geqslant \|2(\theta_0 + \frac{\delta}{2n}b)\| \\
&= (t_1 + t_2)\left\|\left(\frac{t_1}{t_1+t_2}\theta_1 + \frac{t_2}{t_1+t_2}\theta_2\right) + \frac{\delta}{1.5}\left(\frac{t_1}{t_1+t_2}b_1 + \frac{t_2}{t_1+t_2}b_2\right)\right\| \\
&\geqslant (t_1+t_2)\left(\left\|\frac{t_1}{t_1+t_2}\theta_1 + \frac{t_2}{t_1+t_2}\theta_2\right\| - \frac{\delta}{1.5}\left\|\frac{t_1}{t_1+t_2}b_1 + \frac{t_2}{t_1+t_2}b_2\right\|\right) \\
&\geqslant (t_1+t_2)(\delta - \frac{\delta}{1.5}) \\
&\geqslant \frac{t_1\delta}{3}.
\end{aligned}$$

Therefore, $t_1 \leqslant 6(\|\theta_0\|\delta^{-1} + \frac{1}{2n})$. From (4.12),

$$\|x\| \leqslant M + \frac{\delta}{2n} + 6(M\delta^{-1} + \frac{1}{2n})(M + \frac{\delta}{1.5}),$$

where $M := \sup\{\, \|\theta\| \mid \theta \in \Theta \,\} < \infty$. Hence, $U(n,b)$ is bounded.

□

Proposition 4.1.6 *Let $n \in \mathbb{N}$ and $b \in \mathbb{B}^0$. Then for any $x \in S_n$, $P_b^{(n)}(x) = Q_b^{(n)}(x)$.*

Proof

Since $U(n,b) \subseteq \theta_0 + \frac{\delta}{2n}b - S_n$, $Q_b^{(n)}(x) \geqslant P_b^{(n)}(x)$ for any $x \in E$.

For any $x \in S_n$ and $t > 0$, $x + t(\theta_0 + \frac{\delta}{2n}b) \in S_n$ and so

$$x \in t(-\theta_0 - \frac{\delta}{2n}b + S_n). \tag{4.13}$$

From (4.4), $x \in (P_b^{(n)}(x)+\epsilon)(\theta_0+\frac{\delta}{2n}b-S_n)$ for any $\epsilon>0$. This and (4.13) imply that $x \in (P_b^{(n)}(x)+\epsilon)U(n,b)$. Since ϵ is arbitrary, $Q_b^{(n)}(x) \leqslant P_b^{(n)}(x)$ for any $x \in S_n$. This completes the proof

□

By Proposition 4.1.4 and Proposition 4.1.6, we have the following corollary.

Corollary 4.1.1 *For any $n \in \mathbb{N}$ and $b \in \mathbb{B}^0$, $Q_b^{(n)}(x) \leqslant Q_b^{(n)}(x)$ whenever x, $y \in C$ and $x \leqslant_C y$.*

Note that $x+\frac{2n^2}{\delta}\theta_0 = \frac{2n^2}{\delta}\left(\theta_0+\frac{\delta}{2n}(\frac{x}{n})\right) \in \frac{2n^2}{\delta}S_n = S_n$ whenever $x \in A$ and $n > \|A\| := \sup\{\, \|x\| \mid x \in A\,\}$. Thus, by Proposition 4.1.6, $P_b^{(n)}$ and $Q_b^{(n)}$ agree at x. By Theorem 4.1.1 and Theorem 4.1.2, we have the following theorems.

Theorem 4.1.3 (Zheng [21]) *Let A be a bounded subset of E. Then $x \in HE(A,\Theta)$ if and only if $x \in A$ and there exist $b \in \mathbb{B}^0$, $n > \|A\|$ such that*

$$Q_b^{(n)}(x+\frac{2n^2}{\delta}\theta_0) = \inf\{\, Q_b^{(n)}(y+\frac{2n^2}{\delta}\theta_0) \mid y \in A\,\}.$$

Theorem 4.1.4 (Zheng [21]) *Let A be a bounded subset of E and Θ be bounded. Then $x \in SE(A,C)$ if and only if $x \in A$ and there exist $b \in \mathbb{B}^0$, $n > \|A\|$ such that*

$$P_b^{(n)}(x+\frac{2n^2}{\delta}\theta_0) = \inf\{\, P_b^{(n)}(y+\frac{2n^2}{\delta}\theta_0) \mid y \in A\,\}.$$

4.2 Pareto Optimizing and Scalarly Stationary Sequence

In this section, we study the relationship between optimizing and scalarly stationary sequence of vector optimization problem. These results are taken from [4]. Let X and Y be real Banach spaces and $Y^\star$ be the topological dual of Y. $\mathbb{B}$ stands for the unit open ball and $\partial\mathbb{B}$ the unit sphere in any of these spaces. Let

$C \subseteq Y$ be a closed convex pointed partially ordered cone. We also suppose that $int(C) \neq \emptyset$ and $\overline{int(C)} = C$. We note that $0 \notin int(C)$ (for otherwise, C would contain a ball centered at 0 which contradicting the fact that C is pointed).

Let us consider a map $F : X \to Y$,

$$minimize\ F(x)$$
$$s.t.\ x \in X.$$

Let $a \in X$ be a weakly Pareto solution, written as $a \in E_w$, if

$$F(X) - F(a) \subseteq Y \backslash -int(C).$$

Let $a \in X$ be a locally weakly Pareto solution, written as $a \in E_w^{loc}$, if there exists $\epsilon > 0$ such that

$$F(a + \epsilon \mathbb{B}) - F(a) \subseteq Y \backslash -int(C).$$

Remark: $E_w \subseteq E_w^{loc}$ and that, if F is continuous, then E_w is closed in X.

Define

$$C^+ := \{ \lambda \in Y^\star \mid \langle \lambda, y \rangle \geqslant 0 \quad \text{for all } y \in C \},$$
$$C_0^+ := \{ \lambda \in Y^\star \mid \langle \lambda, y \rangle > 0 \quad \text{for all } y \in C \backslash \{ 0 \} \}.$$

Note that $C^+ = \bigcap_{y \in C} \{ \lambda \in Y^\star \mid \langle \lambda, y \rangle \geqslant 0 \}$, so it is closed in $Y^\star$.

Lemma 4.2.1 *$\langle \lambda, y \rangle > 0$, for all $\lambda \in C^+ \backslash \{ 0 \}$ and $y \in int(C)$.*

Proof

Let $\lambda \in C^+ \backslash \{ 0 \}$ and $y \in int(C)$. We have $y + \epsilon \mathbb{B} \subseteq C$ for some $\epsilon > 0$. By the definition of C^+,

$$\langle \lambda, y + \epsilon b \rangle \geqslant 0 \qquad \text{for all } b \in \mathbb{B},$$
$$\langle \lambda, y \rangle \geqslant -\epsilon \langle \lambda, b \rangle \qquad \text{for all } b \in \mathbb{B}.$$

By symmetry of $\mathbb{B}$, we have

$$\langle \lambda, y \rangle \geqslant \epsilon \langle \lambda, b \rangle \qquad \text{for all } b \in \mathbb{B}.$$

$$\text{That is} \quad \langle \lambda, y \rangle \geqslant \epsilon \|\lambda\| > 0 \qquad \text{since } \lambda \neq 0.$$

□

Lemma 4.2.2 *For all $z \in int(C)$, $\inf_{\lambda \in \partial\mathbb{B} \cap C^+} \langle \lambda, z \rangle \geqslant dist(z, Y \backslash C)$.*

Proof

Set $\beta := dist(z, Y \backslash C)$. For any $r \in (0, \beta)$, we have $z + r\mathbb{B} \subseteq C$. Let $\lambda \in \partial\mathbb{B} \cap C^+$. Since $\|\lambda\| = 1$, there exists $e_n \in Y$ with $\|e_n\| = 1$ such that $\langle \lambda, e_n \rangle \to 1$. But $y_n := z - re_n \in z + r\mathbb{B} \subseteq C$, hence $0 \leqslant \langle \lambda, y_n \rangle = \langle \lambda, z \rangle - r\langle \lambda, e_n \rangle$. Let $n \to \infty$, we have $r \leqslant \langle \lambda, z \rangle$. Then let $r \to \beta_-$, we get $\beta \leqslant \langle \lambda, z \rangle$. After taking infimum over $\partial\mathbb{B} \cap C^+$ on the right hand side, the proof is completed.

□

Lemma 4.2.3 *(a) $E_w \supseteq \bigcup_{\lambda \in C^+ \backslash \{0\}} arg \min_{x \in X} \langle \lambda, F(x) \rangle$,*

(b) $E_w^{loc} \supseteq \bigcup_{\lambda \in C^+ \backslash \{0\}} arg \min_{x \in X}^{loc} \langle \lambda, F(x) \rangle$.

Proof

(a)

Suppose $a \notin E_w$. Then $F(x_0) - F(a) \in -int(C)$ for some $x_0 \in X$. By Lemma 4.2.2, we have $\langle \lambda, F(a) - F(x_0) \rangle > 0$ for any $\lambda \in C^+ \backslash \{0\}$ implying that

$$a \notin \bigcup_{\lambda \in C^+ \backslash \{0\}} arg \min_{x \in X} \langle \lambda, F(x) \rangle.$$

By the same trick, we can prove (b).

□

Definition 4.2.1 *F is said to be C-convex if for any $\alpha \in [0,1]$ and for any x, $x' \in X$ we have*

$$F((1-\alpha)x + \alpha x') - (1-\alpha)F(x) - \alpha F(x') \in -C.$$

For convenience, we write $F((1-\alpha)x + \alpha x') \leqslant_C (1-\alpha)F(x) + \alpha F(x')$.

Remark:

(a) $F(\sum_{i=1}^{n} \alpha_i x_i) \leqslant_C \sum_{i=1}^{n} \alpha_i F(x_i)$ for $\alpha_i \geqslant 0$ and $\sum_{i=1}^{n} \alpha_i = 1$.

(b) If F is C-convex then $\lambda \circ F$ is convex for any $\lambda \in C^+$.

Proposition 4.2.1 *Let F be a C-convex map. Then*

$$E_w = \bigcup_{\lambda \in C^+ \backslash \{0\}} arg \min_{x \in X} \langle \lambda, F(x) \rangle = E_w^{loc}.$$

<u>**Proof**</u>

For the first equality, it suffices to prove that $E_w \subseteq \bigcup_{\lambda \in C^+ \backslash \{0\}} arg \min_{x \in X} \langle \lambda, F(x) \rangle$ by Lemma 4.2.3. Let $a \in E_w$. Then

$$F(X) - F(a) \subseteq Y \backslash - int(C). \tag{4.14}$$

We claim that $co(F(X) - F(a)) \cap -int(C) = \emptyset$. Suppose the contrary that $z \in -int(C)$ such that $z = (\sum_{i=1}^{n} \alpha_i F(x_i)) - F(a)$ with $\alpha_i \geqslant 0$ and $\sum_{i=1}^{n} \alpha_i = 1$. By C-convexity of F, $F(\sum_{i=1}^{n} \alpha_i x_i) - F(a) - z \leqslant_C 0$. It follows that

$$F(\sum_{i=1}^{n} \alpha_i x_i) - F(a) \in z - C \subseteq -(int(C) + C) = -int(C),$$

where the last equality holds as $int(C) + C$ is an open set contained in C. It contradicts (4.14), so $co(F(X) - F(a)) \cap -int(C) = \emptyset$. By the separation theorem, there exists $\lambda_0 \in Y^\star \backslash \{0\}$ such that

$$\sup_{c \in C} \langle \lambda_0, -c \rangle \leqslant \inf_{x \in X} \langle \lambda_0, co(F(x) - F(a)) \rangle.$$

Since C is a cone, we must have $\lambda_0 \in C^+ \backslash \{0\}$ by the above inequality (for otherwise, the right-hand side and the left-hand side will go to infinity). Also, by considering $c = 0$ on the left-hand side,

$$0 \leqslant \inf_{x \in X} \langle \lambda_0, co(F(x) - F(a)) \rangle \leqslant \inf_{x \in X} \langle \lambda_0, F(x) - F(a) \rangle.$$

Hence, $a \in arg\min_{x\in X}\langle\lambda_0, F(x)\rangle \subseteq \bigcup_{\lambda\in C^+\backslash\{0\}} arg\min_{x\in X}\langle\lambda, F(x)\rangle$.

The proof of $E_w = E_w^{loc}$ is the same as in the case of a real convex function where any local minimum is a global one.

□

Definition 4.2.2 *Let F be a Frechet differentiable map from X to Y. A point $a \in X$ is called weakly scalarly stationary, written as $a \in S_w$, if there exists $\lambda \in C^+\backslash\{0\}$ such that $\langle\lambda, F'(a)x\rangle = 0$ for all $x \in X$, where $F'(a)$ is the Frechet derivative of F at a.*

Note that $a \in S_w$ if and only if $(C^+\backslash\{0\}) \cap (F'(a)X)^\perp \neq \emptyset$.

Proposition 4.2.2 *Let F be Frechet differentiable and $a \in X$. Then*

$$a \in S_w \quad \textit{if and only if} \quad F'(a)X \cap int(C) = \emptyset.$$

Proof

Suppose that $a \in S_w$ but $x_0 \in F'(a)X \cap int(C) \neq \emptyset$. Then there exists $\lambda_0 \in C^+\backslash\{0\}$ such that λ_0 vanishes on $F'(a)X$ (in particular, $\langle\lambda_0, x_0\rangle = 0$). But this is not possible since Lemma 4.2.1 implies that $\lambda_0 > 0$ on $int(C)$ and hence at x_0.

Conversely, suppose that $F'(a)X \cap int(C) = \emptyset$. Then, by the separation theorem, there exists $\lambda \in Y^\star\backslash\{0\}$ vanishing on $F'(a)X$ such that $\lambda \geqslant 0$ on C. That is, $\lambda \in C^+\backslash\{0\}$ and so $a \in S_w$.

□

Proposition 4.2.3 *Let F be Frechet differentiable. Then $E_w^{loc} \subseteq S_w$. Moreover, if F is C-convex, then $E_w^{loc} = E_w = S_w$.*

Proof

Let $a \in E_w^{loc}$. Note that the set $Y\backslash - int(C)$ is a closed cone (not necessarily convex). For $\epsilon > 0$ small enough and for all $x \in \epsilon\mathbb{B}$,

$$F(a + \alpha x) - F(a) = F'(a)(\alpha x) + o(\alpha x) \in Y\backslash - int(C),$$

for any $\alpha \in (0,1)$. Hence,

$$F'(a)(x) + \frac{o(\alpha x)}{\alpha} \in Y \setminus -int(C),$$

for any $\alpha \in (0,1)$. Let $\alpha \to 0^+$, we have $F'(a)x \in Y\setminus -int(C)$. It follows that $F'(a)X \cap -int(C) = \emptyset$ and so $F'(a)X \cap int(C) = \emptyset$. By Proposition 4.2.2, $a \in S_w$. Suppose that F is C-convex and $x_0 \in S_w$. Then

$$\langle \lambda_0, F'(x_0)x \rangle = \langle (\lambda_0 \circ F)'(x_0), x \rangle = 0$$

for all $x \in X$ and for some $\lambda_0 \in C^+ \setminus \{0\}$. That is, x_0 is a stationary point for the convex function $\lambda_0 \circ F$ (since F is C-convex). We have

$$x_0 \in argmin^{loc}_{x\in X}\langle \lambda_0, F(x)\rangle = arg\min_{x\in X}\langle \lambda_0, F(x)\rangle.$$

Hence, $x_0 \in E^{loc}_w = E_w$ by Lemma 4.2.3 and Proposition 4.2.1.

□

Definition 4.2.3 *(a) The weakly infimal set of F is defined to be*

$$INF_w(F) := \{ y \in \overline{F(X)} \mid F(X) - y \subseteq Y \setminus -int(C) \}.$$

(b) The locally weakly infimal set of F is defined to be

$$INF^{loc}_w(F) := \{ y \in \overline{F(X)} \mid \text{ there exists } \epsilon > 0 \text{ such that } (F(X) \cap (y + \epsilon\mathbb{B})) - y \subseteq Y \setminus -int(C) \}.$$

(c) A sequence $\{x_n\}$ in X is called:

(i) asymptotically weakly Pareto optimizing (a.w.p.) if

$$dist(F(x_n), INF_w(F)) \to 0;$$

(ii) weakly Pareto optimizing (w.p.) if the sequence $\{F(x_n)\}$ converges to an element of $INF_w(F)$;

(iii) weakly scalarly stationary (w.s.s.) if F is Frechet differentiable and there exists a sequence $\{\lambda_n\}$ in $\partial\mathbb{B} \cap C^+$ such that

$$\lim_{n\to\infty} \|\lambda_n \circ F'(x_n)\| = 0. \tag{4.15}$$

(d) For each $\alpha \in Y$, we define the level set by

$$L(\alpha) := \{x \in X \mid F(x) - \alpha \in -C\}.$$

Note that $L(\alpha)$ is closed if F is continuous and $L(\alpha)$ is convex if F is C-convex. In the following, we assume that $INF_w(F)$ is not empty. Then we have the following propositions that characterize the set $INF_w(F)$.

Proposition 4.2.4 *Let F be continuous. Consider for each $\lambda \in C^+\backslash\{0\}$, the set F_λ given by*

$$F_\lambda := \left\{ y \in \overline{F(X)} \mid \langle \lambda, y\rangle = \inf_{x\in X} \langle \lambda, F(x)\rangle \right\}.$$

Then,

$$INF_w(F) \supseteq \bigcup_{\lambda \in C^+\backslash\{0\}} F_\lambda.$$

Moreover, if F is C-convex, then the inclusion becomes an equality.

<u>Proof</u>

By similar arguments in the proofs of Lemma 4.2.3 and Proposition 4.2.1, we can prove the proposition.

□

Proposition 4.2.5 *Assume that F is continuous and let $e \in int(C)$. Then*

$$INF_w(F) = \bigcap_{\eta>0} \bigcup_{\alpha\in \eta e + INF_w(F)} \overline{F(L(\alpha))}.$$

Proof

Let $\alpha_0 \in INF_w(F)$. We can find a sequence $\{x_n\} \subseteq X$ such that $F(x_n) \to \alpha_0$. We claim that for any $\eta > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$F(x_n) \in F(L(\alpha_0 + \eta e)) \quad \text{for all } n \geqslant n_0.$$

Let $z_n := F(x_n) - \alpha_0$. For any $\eta > 0$, since $e \in int(C)$ and $z_n \to 0$, there exists $n_0 \in \mathbb{N}$ such that $e \pm z_n/\eta \in C$ for all $n \geqslant n_0$. Therefore,

$$\begin{aligned} \eta e \pm z_n &\subseteq C && \text{for all } n \geqslant n_0, \\ -\eta e \leqslant_C z_n &\leqslant_C \eta e && \text{for all } n \geqslant n_0, \\ F(x_n) &\leqslant_C \alpha_0 + \eta e && \text{for all } n \geqslant n_0. \end{aligned} \tag{4.16}$$

Note that $F(x_n) \in F(L(F(x_n)))$. From (4.16),

$$F(x_n) \in F(L(F(x_n))) \subseteq F(L(\alpha_0 + \eta e)) \quad \text{for all } n \geqslant n_0.$$

Let $n \to \infty$, we have $\alpha_0 \in \overline{F(L(\alpha_0 + \eta e))}$ for any $\eta > 0$. Hence,

$$\alpha_0 \in \bigcap_{\eta > 0} \bigcup_{\alpha \in \eta e + INF_w(F)} \overline{F(L(\alpha))}.$$

Conversely, let $y \in \bigcap_{\eta>0} \bigcup_{\alpha \in \eta e + INF_w(F)} \overline{F(L(\alpha))}$. If $y \notin INF_w(F)$, then $F(x) - y \in -int(C)$ for some $x \in X$. Let $\eta > 0$ with $F(x) - y + \eta e \in -int(C)$. For this η, there exists $\alpha_0 \in INF_w(F)$ with $y \in \overline{F(L(\eta e + \alpha_0))}$. Hence, $y - \alpha_0 - \eta e \in -C$. Finally, we have $F(x) - y + \eta e + y - \alpha_0 - \eta e \in -int(C) - C = -int(C)$. That is, $F(x) - \alpha_0 \in -int(C)$ which contradicts with the choice of α_0. This completes the proof.

□

For any $\rho > 0$ and $x \in X$, define

$$\eta(\rho, x) := 1 + \sup\{\frac{dist\,(x', L(F(x)))}{\|F(x) - F(x')\|} \mid \|F(x) - F(x')\| \geqslant \rho \quad \text{and } F(x) - F(x') \in -int(C)\}$$

and $\eta(\rho) := \sup_{x \in X} \eta(\rho, x)$ with the convention $\sup \emptyset = 0$.

Note that

$$dist\left(x', L(F(x))\right) \leqslant \eta(\rho, x)\|F(x) - F(x')\| \leqslant \eta(\rho)\|F(x) - F(x')\|$$

whenever $\|F(x) - F(x')\| \geqslant \rho$ and $F(x) - F(x') \in -int(C)$.

Theorem 4.2.1 (Bernoussi, Bolintineanu and Chou [4]**)** *Let X be a reflexive Banach space, F be Frechet differentiable and C-convex. Suppose $\{ x_n \}$ is a w.s.s. sequence. Then we have the followings:*

(a) If $\{ x_n \}$ converges, then its limit is a weakly Pareto point.

(b) Assume that $\eta(\rho, x) < +\infty$ for all $\rho > 0$ and for all $x \in X$. If the sequence $\{ F(x_n) \}$ converges, then $\{ x_n \}$ is a w.p. sequence.

(c) Assume that $0 < \eta(\rho) < +\infty$ for all $\rho > 0$. If there exists a sequence $\{ y_n \}$ in X such that

1. $F(y_n) - F(x_n) \in -int(C)$ and $dist(F(y_n), INF_w(F)) \to 0$.

2. there exists $K > 0$ such that for all $n \in \mathbb{N}$,

$$\langle \lambda_n, F(x_n) - F(y_n) \rangle \geqslant K\|F(x_n) - F(y_n)\|,$$

where $\{ \lambda_n \} \subseteq \partial\mathbb{B} \cap C^+$ is a sequence verifying (4.15)*,*

then $\{ x_n \}$ is an a.w.p. sequence.

<u>Proof</u>

<u>(a)</u>

Let $x_n \to x$ and $\{ \lambda_n \}$ be a sequence verifying (4.15). Suppose the contrary that $x \notin E_w$. Then there exists $b \in X$ such that $F(b) - F(x) \in -int(C)$. Therefore, we can find an $n_0 \in \mathbb{N}$ with $F(b) - F(x_n) \in -int(C)$ for all $n \geqslant n_0$. Hence, for any $n \geqslant n_0$,

$$\begin{aligned}
0 &< dist(F(x) - F(b), Y\backslash C) \\
&\leqslant \|F(x) - F(x_n)\| + dist(F(x_n) - F(b), Y\backslash C) \\
&\leqslant \|F(x) - F(x_n)\| + \langle \lambda_n, F(x_n) - F(b)\rangle && \text{by Lemma4.2.2} \\
&\leqslant \|F(x) - F(x_n)\| + \langle \lambda_n, F^{'}(x_n)(x_n - b)\rangle && \text{by the convexity of } \lambda_n \circ F \\
&\leqslant \|F(x) - F(x_n)\| + \|\lambda_n \circ F^{'}(x_n)\|\|x_n - b\|.
\end{aligned}$$

Since $\|F(x) - F(x_n)\| \to 0$ and $\|\lambda_n \circ F^{'}(x_n)\| \to 0$, it follows that

$$0 < dist(F(x) - F(b), Y\backslash C) \leqslant \|F(x) - F(x_n)\| + \|\lambda_n \circ F^{'}(x_n)\|\|x_n - b\| \to 0$$

as n $\to \infty$. It is impossible.

(b)

Let $F(x_n) \to y$. Then $y \in \overline{F(X)}$. If $y \notin INF_w(F)$: there exists $x \in X$ such that $F(x) - y \in -int(C)$, then there exists $\rho > 0$ such that $F(x) - y + 2\rho\mathbb{B} \subseteq -int(C)$. So $F(x) - F(x_n) \in F(x) - y + \rho\mathbb{B}$ for n large enough, say $n \geqslant n_0$. Since $0 \notin -int(C)$, it follows that $0 \neq F(x) - F(x_n) + \rho\mathbb{B} \subseteq F(x) - y + 2\rho\mathbb{B} \subseteq -int(C)$ for all $n \geqslant n_0$. Therefore,

$$dist(F(x_n) - F(x), Y\backslash C) \geqslant \rho \text{ and } \|F(x) - F(x_n)\| \geqslant \rho > 0$$

for all $n \geqslant n_0$. By the definition of $\eta(\rho, x)$, we have

$$dist\,(x_n, L(F(x))) \leqslant \eta(\rho, x)\|F(x) - F(x_n)\| \quad \text{for all } n \geqslant n_0. \tag{4.17}$$

Let z_n be the projection of x_n on the closed convex set $L(F(x))$. For all $n \geqslant n_0$, $F(x) \geqslant_C F(z_n) \geqslant_C F(x_n) + F^{'}(x_n)(z_n - x_n)$. Hence, by Lemma 4.2.2 and (4.17),

we have

$$\begin{aligned}
\|\lambda_n \circ F^{'}(x_n)\| \cdot \eta(\rho, x) \cdot \|F(x) - F(x_n)\| &\geqslant \|\lambda_n \circ F^{'}(x_n)\| \|x_n - z_n\| \\
&\geqslant \langle \lambda_n, F^{'}(x_n)(x_n - z_n) \rangle \\
&\geqslant \langle \lambda_n, F(x_n) - F(x) \rangle \\
&\geqslant dist(F(x_n) - F(x), Y \backslash C) \geqslant \rho.
\end{aligned}$$

This is impossible, because $\|\lambda_n \circ F^{'}(x_n)\| \to 0$ and $\|F(x) - F(x_n)\| \to \|F(x) - y\|$.

(c)

Suppose that $\{ x_n \}$ is not a.w.p.. Then we can find an $\rho > 0$ with

$$dist(F(x_n), INF_w(F)) > 2\rho \quad \text{for infinitely many } n.$$

Consider subsequence if necessary, we may assume that it holds for all n. By assumption (1) and the above inequality, there exists $n_0 \in \mathbb{N}$ such that

$$\|F(x_n) - F(y_n)\| \geqslant \rho \quad \text{for all } n \geqslant n_0. \tag{4.18}$$

Let z_n be the projection of x_n on the closed convex set $L(F(y_n))$. Then $\|x_n - z_n\| = dist(x_n, L(F(y_n)))$ and

$$F(z_n) - F(y_n) \in -C. \tag{4.19}$$

We have

$$F(y_n) \geqslant_C F(z_n) \geqslant_C F(x_n) + F^{'}(x_n)(z_n - x_n).$$

Hence,

$$\begin{aligned}
\|\lambda_n \circ F^{'}(x_n)\| \|x_n - z_n\| &\geqslant \langle \lambda_n, F^{'}(x_n)(x_n - z_n) \rangle \\
&\geqslant \langle \lambda_n, F(x_n) - F(y_n) \rangle \\
&\geqslant K \|F(x_n) - F(y_n)\| \quad \text{by assumption (2)} \\
&\geqslant \frac{K}{\eta(\rho)} dist\,(x_n, L(F(y_n))) \\
&\geqslant \frac{K}{\eta(\rho)} \|x_n - z_n\|.
\end{aligned}$$

Since $\|\lambda_n \circ F'(x_n)\| \to 0$ as $n \to 0$, we must have $x_n = z_n$ for n large enough. Combining this result with assumption (1), (4.18) and (4.19), we have C is not pointed. This is impossible.

$\square$

Bibliography

[1] A. Auslender and J. P. Crouzeix, *Well behaved asymptotical convex functions*, Analyse Lineare, (1989), 101-112.

[2] A. Auslender, R. Cominetti and J. P. Crouzeix, *Convex functions with unbounded level sets and applications to duality theory*, SIAM Journal on Optimization, 3 (1993), 669-687.

[3] V. M. Alekseev, V. M. Tikhomirov and S. V. Fomin, *Optimal control*, Nauka, Moscow, 1979.

[4] B. Bernoussi, S. Bolintineau and C. C. Chou, *Pareto optimizing and scalarly stationary sequences*, Journal of Analysis and Applications, Vol. 220, No. 2 (1998), 553-561.

[5] J. M. Borwein and D. Zhuang, *Super efficiency in vector optimization*, Transactions of the American Mathematical Society, Vol. 338 (1993), 105-122.

[6] F. H. Clarke, *Optimization and nonsmooth analysis*, Centre de recherches mathematiques, 1989.

[7] Sien Deng, *Computable error bounds for convex inequality systems in reflexive Banach spaces*, SIAM Journal on Optimization, Vol. 7, No. 1 (1997), 274-179.

[8] J. P. Hiriart-Urruty and C. Lemarechal, *Convex analysis and minimization algorithms I*, Springer-Verlag, Berlin, 1993.

[9] J. P. Hiriart-Urruty and C. Lemarechal, *Convex analysis and minimization algorithms II*, Springer-Verlag, Berlin, 1993.

[10] Alan J. Hoffman, *On approximate solutions of systems of linear inequalities*, Journal of Research of the National Bureau of Standards, Vol. 49, No. 4 (1952), 263-265.

[11] L. Hormander, *On the division of distributions by polynomials*, Ark. Mat., 3 (1958), 555-568.

[12] A. D. Ioffe and V. M. Tikhomirov, *Theory of extremal problems*, Nauka, Moscow, 1974.

[13] Adrian S. Lewis and J. S. Pang, *Error bounds for convex inequality systems*, preprint.

[14] Xiao Dong Luo and Zhi Quan Luo, *Extension of Hoffman's error bound to polynomial systems*, SIAM Journal on Optimization, Vol. 4, No. 2 (1994), 383-392.

[15] O. L. Mangasarian, *A condition number for differentiable convex inequalities*, Math. Oper. Res., 10 (1985), 175-179.

[16] O. L. Mangasarian, *Error bounds for nondifferentiable convex inequalities under a strong Slater constraint qualification*, Mathematical Programming, Vol. 83 (1998), 187-194.

[17] O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969 (reprint from SIAM Classic in Applied Mathematics 10, 1994).

[18] R. R. Phelps, *Convex functions, monotone operators and differentiability*, Lecture Notes in Mathematics, No. 1364, Springer-verlag, (1989).

[19] R. T. Rockafellar, *Convex analysis*, Princeton University Press, 1970.

[20] R. T. Rockafellar, *The theory of subgradients and its applications to problems of optimization. Convex and nonconvex functions*, Heldermann Verlag Berlin, 1981.

[21] Xi Yin Zheng, *Scalarization of Henig proper efficient points in a normed space*, preprint.